June 1961

BIOMETRICS

Vol. 17 No. 2

JOURNAL OF THE BIOMETRIC SOCIETY

Articles

On Additivity in the Analysis of Variance — *R. C. Elston*

A Note on Some Growth Patterns in a Simple Theoretical Organism — *J. A. Nelder*

The Partial Diallel Cross — *O. Kempthorne and R. N. Curnow*

Sensory Testing by Triple Comparisons — *G. T. Park*

Rapid Chi-Square Test of Significance for Three-Part Ratios — *Charles E. Gates and Benjamin H. Beard*

Generating Unbiased Ratio and Regression Estimators — *W. H. Williams*

A Biometric Theory of Middle and Long Distance Track Records — *Malcolm E. Turner and Eleanor D. Campbell*

Small Sample Behavior of Slope Estimators in a Linear Functional Relation — *Martin Dorff and John Gurland*

Statistics for a Diagnostic Model — *Adrianus J. van Woerkom and Keeve Brodman*

Queries and Notes

Three-Quarter Replicates of 2^4 and 2^5 Designs — *Peter W. M. John*

On the Extension of Stevens' Tables for Asymptotic Regression — *S. Lipton*

Corrected Error Rates for Duncan's New Multiple Range Test — *H. Leon Harter*

Multiple Comparisons between Treatments and a Control — *C. W. Dunnett*

Error Rates in Multiple Comparisons — *R. G. D. Steel*

Book Reviews

D. S. Falconer: Introduction to Quantitative Genetics — *W. F. Bodmer*

S. Goldberg: Probability: An Introduction — *F. N. David*

THE BIOMETRIC SOCIETY

The Biometric Society is an international society devoted to the mathematical and statistical aspects of biology. Biologists, mathematicians, statisticians, and others interested in its objectives are invited to become members. Through its regional organizations the Society sponsors regional and local meetings. National secretaries serve the interests of members in Denmark, India, Japan, the Netherlands, Norway, Sweden, and Switzerland, and there are many members at large.

Biometrics, the journal of the Biometric Society, is published quarterly by the William Byrd Press, Inc., 1407 Sherwood Ave., Richmond 22, Va., for the Biometric Society. Its general objects are to promote and to extend the use of mathematical and statistical methods in pure and applied biological sciences, by describing and exemplifying developments in these methods and their applications in a form readily assimilable by experimental scientists. It is also intended to provide a medium for exchange of ideas by experimenters and those concerned primarily with analysis and the development of statistical methodology.

Original papers, and authoritative expository or review articles or critiques, will be accepted for publication in *Biometrics* if judged consistent with these general aims. Predominantly analytical or methodological papers should contribute specifically to the formulation of quantitative hypotheses, to the interpretation of data, or to the planning or analysis of experiments or surveys. Papers dealing with biological subjects should report conclusions of definite applicability reached by mathematical or statistical analysis, so described as to facilitate possible use of the procedure in other fields of biology or related sciences. Authors of statistical and mathematical papers are asked specifically to prepare introductions and examples to clearly indicate the natures of the problems in biology and related fields to which they apply.

Technical notes or problems for consideration under the heading of Queries and Notes are invited.

Information for contributors is given on the inside back cover.

ANNUAL DUES AND MEMBERSHIPS FOR 1961

Memberships (including dues and subscriptions to this journal)	
Membership (U.S.A. and Canada)	$ 7.00
Student membership (U.S.A. and Canada)	4.00
Membership for others	4.50
Student membership for others	2.75
Sustaining membership (including two subscriptions to this journal)	100.00
Subscriptions	
Non-members of the Biometric Society	7.00

All dues, memberships, and subscriptions are payable in U.S.A. currency. Information concerning the Society and memberships may be obtained from its Secretary, M. J. R. Healy, Statistics Department, Rothamsted Experimental Station, Harpenden, Herts., England. Non-member subscriptions are payable to the Managing Editor, *Biometrics*, Department of Statistics, The Florida State University, Tallahassee, Florida, U.S.A.

BUSINESS OFFICE OF *BIOMETRICS*: Department of Statistics, The Florida State University, Tallahassee, Florida, U.S.A. Non-member subscriptions, changes of address for non-member subscriptions, and undeliverable copies should be sent to this office.

BUSINESS OFFICE OF THE SOCIETY. 110 Virginia Road, Oak Ridge, Tennessee, U.S.A. Members of the Biometric Society are advised to send changes of address and all correspondence regarding member subscriptions to their Regional or National Secretary or, for members-at-large to the Biometric Society, P.O. Box 2017, Oak Ridge, Tennessee, U.S.A.

Second-class postage paid at Richmond, Virginia. *Biometrics* is published quarterly—in March, June, September, and December.

The Biometric Society

BIOMETRICS

FOUNDED BY THE BIOMETRICS SECTION OF THE AMERICAN STATISTICAL ASSOCIATION

TABLE OF CONTENTS

Articles

On Additivity in the Analysis of Variance ..... R. C. Elston 209

A Note on Some Growth Patterns in a Simple Theoretical Organism
J. A. Nelder 220

The Partial Dialled Cross ..... O. Kempthorne and R. N. Curnow 229

Sensory Testing by Triple Comparisons ..... G. T. Park 251

Rapid Chi-Square Test of Significance for Three-Part Ratios
Charles E. Gates and Benjamin H. Beard 261

Generating Unbaised Ratio and Regression Estimators ..... W. H. Williams 267

A Biometric Theory of Middle and Long Distance Track Records
Malcolm E. Turner and Eleanor D. Campbell 275

Small Sample Behavior of Slope Estimators in a Linear Functional Relation
Martin Dorff and John Gurland 283

Statistics for a Diagnostic Model
Adrianus J. van Woerkom and Keeve Brodman 299

Queries and Notes

Three-Quarter Replicates of 2^4 and 2^5 Designs ..... Peter W. M. John 319

On the Extention of Stevens' Tables for Asymptotic Regression .. S. Lipton 321

Corrected Error Rates for Duncan's New Multiple Range Test
H. Leon Harter 321

Multiple Comparisons between Treatments and a Control .. C. W. Dunnett 324

Error Rates in Multiple Comparisons ..... R. G. D. Steel 326

Book Reviews

D. S. Falconer: Introduction to Quantitative Genetics .... W. F. Bodmer 329

S. Goldberg: Probability: An Introduction ..... F. N. David 330

Abstracts 331

Correction 333

The Biometric Society 334

News and Announcements 340

Number 2 June 1961 Volume 17

BIOMETRICS

Editor

Ralph A. Bradley

Editorial Borad

Editorial Associates and Committee Members: C. I. Bliss, Irwin Bross, E. A. Cornish, S. Lee Crump, H. A. David, W. J. Dixon, Mary Elveback, D. J. Finney, W. A. Glenn, J. W. Hopkins, O. Kempthorne, Leopold Martin, Horace W. Norton, S. C. Pearce, J. G. Skellam, and Georges Teissier. *Managing Editor*: Ralph A. Bradley.

Former Editors

Gertrude M. Cox—*Founding Editor*

John W. Hopkins—*Past Editor*

Officers of the Biometric Society

General Officers

President: Leopold Martin

Secretary: M. J. R. Healy; *Treasurer*: M. A. Kastenbaum

Council: M. S. Bartlett, W. U. Behrens, C. I. Bliss, A. Buzzati-Traverso, D. G. Chapman, W. J. Dixon, C. W. Emmens, D. J. Finney, C. G. Fraga, Jr., Anne Lenger, C. C. Li, A. Linder, K. Mather, C. R. Rao, P. V. Sukhatme, G. Teissier, J. W. Tukey, G. S. Watson, F. Yates.

Regional Officers

Region	*President*	*Secretary*	*Treasurer*
Australasian	M. Belz	W. B. Hall	G. W. Rogerson
Belgian	M. Welsch	L. Martin	P. Gilbert
Brazilian	A. M. Penha	A. Conagin	A. Groszmann
British	J. A. Fraser Roberts	C. D. Kemp	P. A. Young
E. N. American	Oscar Kempthorne (Henry L. Lucas)	Erwin L. LeClerg	Donald A. Gardiner
French	Ph. L'Heritier	Sully Ledermann	Sully Ledermann
German	O. Heinisch	R. Wette	M. P. Geppert
Italian	G. Montalenti	R. Scossiroli	F. Sella
W. N. American	W. A. Taylor	W. A. Becker	Bernice Brown

National Secretaries

Denmark	N. F. Gjeddebæk	*Netherlands*	H. de Jonge
India	A. R. Roy	*Norway*	L. K. Strand
Japan	M. Hatamura	*Sweden*	H. A. O. Wold
		Switzerland	H. L. LeRoy

ON ADDITIVITY IN THE ANALYSIS OF VARIANCE[1]

R. C. ELSTON
Department of Biostatistics, University of North Carolina
Chapel Hill, North Carolina, U. S. A.

INTRODUCTION

The problem relating to additivity in the analysis of variance is twofold. In the first place we wish to know whether we can remove any of the non-additivity present in our data, and in the second place we wish to know, given that it can be done, how to do so. Here two methods that have already been proposed for testing for non-additivity are developed from a somewhat different viewpoint, in an attempt to clarify their properties; and a generalization is given of a method that has been suggested for finding an appropriate transformation of the data.

TESTING FOR REMOVABLE NON-ADDITIVITY

Definition. If μ_{ij} is the subclass mean corresponding to the i-th level of A and to the j-th level of B in any 2-way classification, then we shall say that A and B are additive (or the μ_{ij} show additivity) if and only if there exist constants α_i , β_j such that

$$\mu_{ij} = \alpha_i + \beta_j , \qquad \text{all} \quad i, j.$$

For an n-way classification, denote the subclass means by μ_q , where q is an n-tuple of numbers designating the particular subclass; then there is additivity if and only if there exist constants α_{i_1} , β_{i_2} , $\cdots$, ν_{i_n} such that

$$\mu_q = \mu_{i_1 i_2 \cdots i_n} = \alpha_{i_1} + \beta_{i_2} + \cdots + \nu_{i_n} , \qquad \text{all} \quad i_1 , i_2 , \cdots , i_n .$$

Now whether or not the μ_q show additivity depends on the scale on which they are measured. In practice we measure our data on any scale that happens to be convenient, and from these measurements estimate the μ_q ; and, unless we transform the data, we are estimating the μ_q as measured on the same scale as that on which the data are measured. But, in many cases, the estimates we obtain are of far greater value to us if we can assume that the μ_q show additivity. Thus,

[1]Supported in part by departmental fellowship funds provided by U. S. Public Health Service.

if we analyze a set of data y_q and find we have to reject the hypothesis that the corresponding μ_q show additivity, we should try and find a transformation of y_q to x_q such that the $E(x_q)$ show additivity (where E denotes expectation). Since the y_q are measured on a continuous scale, it is reasonable to consider only continuous functions of the y_q , and for such a transformation to be of any practical use it must be a *single-valued* function. With the additional reasonable requirement that the inverse function also be single-valued, this is equivalent to requiring the function (or its inverse) to be strictly monotonic.

From here on the symbol μ_q will be used to denote subclass means measured on a scale on which they show additivity, and we shall denote by y_q the data measured on any convenient scale. Suppose there exists a scale such that when the data are measured on it, and we denote these measurements by x_q ,

$$E(x_q) = \mu + \mu_q \quad \text{(where } \mu \text{ is any constant)}. \tag{1}$$

Then the $E(x_q)$ show additivity. Furthermore, there is no loss of generality when we assume

$$\max \mu_q + \min \mu_q = 0, \qquad \max \mu_q \geq 0, \qquad \min \mu_q \leq 0. \tag{2}$$

(We assume μ_q lies within a definite range, determined by the extent of our data.)

If $E(y_q)$ is a linear function of $E(x_q)$, then the $E(y_q)$ show additivity. If $E(y_q)$ is a non-linear function of $E(x_q)$, then we can approximate this function by a quadratic polynomial, i.e.

$$E(y_q) \fallingdotseq a_0 + a_1 E(x_q) + a_2 [E(x_q)]^2. \tag{3}$$

If $E(y_q)$ is better approximated by such a quadratic polynomial than by any linear polynomial in $E(x_q)$, then, provided our data lie within the range where this quadratic polynomial is strictly monotonic, there exists a single-valued function such that, when applied to the data, the expectations of the transformed data come nearer to satisfying the condition for additivity.

Assuming exact equality in (3) and using (1) we may rewrite it as a polynomial in μ_q , say

$$E(y_q) = a + b\mu_q + c\mu_q^2 . \tag{4}$$

Differentiating with respect to μ_q we obtain $b + 2c\mu_q$, and here there is a maximum or minimum. Thus $E(y_q)$ will be strictly monotonic over the two ranges

(i) $\quad b + 2c \min \mu_q > 0$

and

(ii) $\quad b + 2c \max \mu_q < 0.$

Now using (2) it follows that if $\max \mu_a - \min \mu_a < b/|c|$, (i) is satisfied; and if $\max \mu_a - \min \mu_a < -b/|c|$, (ii) is satisfied (and in this case we must have $b < 0$).

It is clear from this that there is no point in considering the case $b = 0$, and so we can arbitrarily let $b = 1$ (i.e. divide our model through by b) and rewrite (4):

$$E(y_a) = \mu + \mu_a + c\mu_a^2 ,$$

where μ and c are appropriately redefined; and this is a strictly monotonic function over the range

$$\max \mu_a - \min \mu_a < \frac{1}{|c|}. \tag{5}$$

What we wish to do, then, is to assume a model of this form with the y_a normally and independently distributed, and test the null hypothesis that $c = 0$. If we reject this hypothesis, then, provided (5) holds, there exists a monotonic function of the data (whose inverse is a quadratic polynomial) that will bring us closer to additivity. In order to test this hypothesis we should like to find the reduction in sum of squares $R(\mathrm{I})$ due to fitting the model under the null hypothesis, and the reduction, $R(\mathrm{II})$, due to fitting the full model. These reductions are given by

$$R(\mathrm{I}) = \sum y_a^2 - \min \sum (y_a - \mu - \mu_a)^2,$$

and

$$R(\mathrm{II}) = \sum y_a^2 - \min \sum (y_a - \mu - \mu_a - c\mu_a^2)^2,$$

where the summations are over all observations, and μ, μ_a and c are estimated from the data. We can find $R(\mathrm{I})$ in the usual way; but when we set up the normal equations to find $R(\mathrm{II})$, we obtain a set of cubic equations that are not easy to solve.

The problem can be simplified, and an approximate solution obtained, as follows. For the sake of clarity consider the 2-way classification, one observation per subclass. Under the null hypothesis $c = 0$ we have

$$(\mathrm{I}) \quad E(y_{ij}) = \mu + \alpha_i + \beta_j ,$$

and otherwise

$$(\mathrm{II}) \quad E(y_{ij}) = \mu + \alpha_i + \beta_j + c(\alpha_i + \beta_j)^2.$$

Now the only extra term in the right hand side of (II) that depends on both i and j is the term $2c\alpha_i\beta_j$, and it is only if the absolute magni-

tude of this term is ever large that (II) can lead to a significantly smaller residual sum of squares than (I). We may therefore (neglecting the coefficient 2) consider the approximately equivalent model, (II′), given by

$$E(y_{ij}) = \mu + \alpha_i + \beta_j + c\alpha_i\beta_j \ ; \tag{6}$$

i.e. we assume $\alpha_i \fallingdotseq \alpha_i + c\alpha_i^2$ and $\beta_j \fallingdotseq \beta_j + c\beta_j^2$, and this is reasonable when we remember that c must satisfy (5) for monotonicity.

In order to obtain R(II′), the reduction in sum of squares due to fitting the model (6), it is necessary to minimize $\sum_{ij} (y_{ij} - \mu - \alpha_i - \beta_j - c\alpha_i\beta_j)^2$, which results in the following normal equations:

$$\sum_{ij} (\hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j + \hat{c}\hat{\alpha}_i\hat{\beta}_j) = \sum_{ij} y_{ij} \tag{7}$$

$$\sum_{j} (\hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j + \hat{c}\hat{\alpha}_i\hat{\beta}_j)(1 + \hat{c}\hat{\beta}_j) = \sum_{j} y_{ij}(1 + \hat{c}\hat{\beta}_j)$$

$$\sum_{i} (\hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j + \hat{c}\hat{\alpha}_i\hat{\beta}_j)(1 + \hat{c}\hat{\alpha}_i) = \sum_{i} y_{ij}(1 + \hat{c}\hat{\alpha}_i)$$

$$\sum_{ij} (\hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j + \hat{c}\hat{\alpha}_i\hat{\beta}_j)(\hat{\alpha}_i\hat{\beta}_j) = \sum_{ij} y_{ij}\hat{\alpha}_i\hat{\beta}_j \ ,$$

where $\hat{\mu}$, $\hat{\alpha}_i$, $\hat{\beta}_j$ are a set of values for μ, α_i , β_j that minimize the above sum of squares. If we let $\sum_i \hat{\alpha}_i = \sum_j \hat{\beta}_j = 0$, we obtain

$$\sum_{ij} \hat{\mu} = \sum_{ij} y_{ij} \ , \tag{8}$$

$$\hat{\alpha}_i = \sum_{j} (y_{ij} - \hat{\mu} - \hat{\beta}_j)(1 + \hat{c}\hat{\beta}_j) / \sum_{j} (1 + \hat{c}\hat{\beta}_j)^2, \tag{9}$$

$$\hat{\beta}_j = \sum_{i} (y_{ij} - \hat{\mu} - \hat{\alpha}_i)(1 + \hat{c}\hat{\alpha}_i) / \sum_{i} (1 + \hat{c}\hat{\alpha}_i)^2, \tag{10}$$

$$\hat{c} = \sum_{ij} y_{ij}\hat{\alpha}_i\hat{\beta}_j / \sum_{ij} (\hat{\alpha}_i\hat{\beta}_j)^2. \tag{11}$$

We can obtain $\hat{\mu}$ directly from (8), but $\hat{\alpha}_i$, $\hat{\beta}_j$ and $\hat{c}$ have to be obtained by iteration. We first calculate $\hat{\alpha}_i$ and $\hat{\beta}_j$ from (9) and (10) assuming $\hat{c} = 0$, and then, using these values of $\hat{\alpha}_i$ and $\hat{\beta}_j$, we calculate ${}_1\hat{c}$, a first approximation to $\hat{c}$, from (11). Then assuming $\hat{c} = {}_1\hat{c}$ we recalculate $\hat{\alpha}_i$ and $\hat{\beta}_j$ from (9) and (10), and use these new values to obtain a second approximation to $\hat{c}$ from (11). The process is continued until all the estimates become stable. R(II′) is then given by

$$\sum_{ij} y_{ij}^2 - \sum_{ij} (y_{ij} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j - \hat{c}\hat{\alpha}_i\hat{\beta}_j)^2$$

$$= \hat{\mu} \sum_{ij} y_{ij} + \sum_{i} \hat{\alpha}_i \sum_{j} y_{ij} + \sum_{j} \hat{\beta}_j \sum_{i} y_{ij} + \hat{c} \sum_{ij} y_{ij}\hat{\alpha}_i\hat{\beta}_j$$

when $\sum_i \hat{\alpha}_i = \sum_j \hat{\beta}_j = 0$ and the estimates are such that this reduction is a maximum.

This method of solution was given by Ward and Dick [1952], but they considered the model (6) for a different reason; they wished to test the appropriateness of a multiplicative model of the form

$$E(y_{ij}) = (\mu + \alpha_i)(\mu' + \beta_j) = \mu\mu' + \mu'\alpha_i + \mu\beta_j + \alpha_i\beta_j ,$$

and this led them to assume the model (6). They state that the F-test for c is only approximate. If we assume as our full model

$$y_{ij} = \mu + \alpha_i + \beta_j + c\alpha_i\beta_j + \epsilon_{ij} ,$$

where the ϵ_{ij} are independent and $N(0, \sigma^2)$, then to test the null hypothesis that $c = 0$ we could take as our statistic

$$F = \frac{R(\text{II}') - R(\text{I})}{[\sum_{ij} y_{ij}^2 - R(\text{II}')]/(n-1)} ,$$

where there would have been n degrees of freedom associated with the residual sum of squares if the model did not include the term $c\alpha_i\beta_j$. But whether or not this statistic actually does follow the F-distribution on 1 and $(n-1)$ degrees of freedom under the null hypothesis requires further investigation.

Now consider the sum of squares due to ${}_1\hat{c}$, the first approximation to $\hat{c}$, which is given by

$$\begin{aligned} R(\text{II}'') - R(\text{I}) &= \sum_{ij} y_{ij}^2 - \sum_{ij} (y_{ij} - \tilde{\mu} - \tilde{\alpha}_i - \tilde{\beta}_j - {}_1\hat{c}\tilde{\alpha}_i\tilde{\beta}_j)^2 \\ &\quad - [\sum_{ij} y_{ij}^2 - \sum_{ij} (y_{ij} - \tilde{\mu} - \tilde{\alpha}_i\tilde{\beta}_j)^2] \\ &= -{}_1\hat{c} \sum_{ij} [{}_1\hat{c}(\tilde{\alpha}_i\tilde{\beta}_j)^2 - 2(\tilde{\alpha}_i\tilde{\beta}_j y_{ij} - \tilde{\mu}\tilde{\alpha}_i\tilde{\beta}_j - \tilde{\alpha}_i^2\tilde{\beta}_j - \tilde{\alpha}_i\tilde{\beta}_j^2)] \qquad (12) \end{aligned}$$

where $\tilde{\mu}$, $\tilde{\alpha}_i$ and $\tilde{\beta}_j$ are values of $\hat{\mu}$, $\hat{\alpha}_i$ and $\hat{\beta}_j$ satisfying (7), (9) and (10) when $\hat{c} = 0$, i.e. we have

$$y_{ij} = \tilde{\mu} + \tilde{\alpha}_i + \tilde{\beta}_j + \tilde{\epsilon}_{ij} \qquad (13)$$

where $\sum_{ij} \tilde{\epsilon}_{ij}^2$ is a minimum. Now from (11) we have ${}_1\hat{c} = \sum_{ij} y_{ij}\tilde{\alpha}_i\tilde{\beta}_j / \sum_{ij} (\tilde{\alpha}_i\tilde{\beta}_j)^2$. Using this and (13), and letting $\sum_i \tilde{\alpha}_i = \sum_j \tilde{\beta}_j = 0$, we find from (12),

$$R(\text{II}'') - R(\text{I}) = \frac{[\sum_{ij} \tilde{\alpha}_i\tilde{\beta}_j y_{ij}]^2}{\sum_{ij} (\tilde{\alpha}_i\tilde{\beta}_j)^2} = \frac{[\sum_{ij} \tilde{\alpha}_i\tilde{\beta}_j\tilde{\epsilon}_{ij}]^2}{\sum_{ij} (\tilde{\alpha}_i\tilde{\beta}_j)^2} .$$

From this it is seen that the sum of squares due to ${}_1\hat{c}$ is just the sum of squares proposed by Tukey [1949a] to test for non-additivity; the general formula given by Tukey [1955] is easily shown to be identical

with the above for a two-way classification when $\sum_i \tilde{\alpha}_i = \sum_j \tilde{\beta}_j = 0$. Thus in Tukey's test for non-additivity we assume the model (II″)

$$y_{ij} = \mu + \alpha_i + \beta_j + c[\alpha_i\beta_j]_{ij} + \epsilon_{ij} ,$$

ϵ_{ij} independent, $N(0, \sigma^2)$, where $c[\alpha_i\beta_j]_{ij}$ indicates that the term is part of the interaction. This is to be interpreted as follows. We let

$$y_{ij} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij}$$

where $\sum_i (\alpha\beta)_{ij} = 0$ for all j and $\sum_j (\alpha\beta)_{ij} = 0$ for all i. This uniquely defines $\mu + \alpha_i + \beta_j$ as population parameters, but not α_i or β_j . Now for any set of α_i and β_j that are consistent with this definition let

$$\alpha_i\beta_j = \alpha_i' + \beta_j' + (\alpha\beta)_{ij}'$$

where $\sum_i (\alpha\beta)_{ij}' = 0$ for all j and $\sum_j (\alpha\beta)_{ij}' = 0$ for all i. Then this uniquely defines $(\alpha\beta)_{ij}'$, and this is what is represented above by $[\alpha_i\beta_j]_{ij}$. The parameter c in model (II″) can be thought of as the regression of the interaction effect $(\alpha\beta)_{ij}$ on the product of the main effects. We then test the null hypothesis that $c = 0$ using the statistic

$$F = \frac{R(\text{II}'') - R(\text{I})}{[\sum_{ij} y_{ij}^2 - R(\text{II}'')]/(n-1)}. \tag{14}$$

Now since the y_{ij} are normally and independently distributed with variance σ^2, $R(\text{II}'') - R(\text{I})$ is distributed, for fixed $\tilde{\alpha}_i$ and $\tilde{\beta}_j$, as $\sigma^2\chi^2$ on 1 degree of freedom; and under the null hypothesis this is a central χ^2-distribution, for then $E(\tilde{\epsilon}_{ij}) = 0$. Also, since the distribution is the same for all fixed $\tilde{\alpha}_i$ and $\tilde{\beta}_j$, this is in fact the marginal distribution of $R(\text{II}'') - R(\text{I})$ as obtained from the joint distribution of $\tilde{\alpha}_i$, $\tilde{\beta}_j$ and $R(\text{II}'') - R(\text{I})$, as was pointed out by Tukey. Furthermore, $\sum_{ij} y_{ij}^2 - R(\text{II}'') = \sum_{ij} \tilde{\epsilon}_{ij}^2 - [R(\text{II}'') - R(\text{I})]$ can be shown to be independently distributed as $\sigma^2\chi^2$ on $(n-1)$ degrees of freedom, n being the number of degrees of freedom associated with $\sum_{ij} \tilde{\epsilon}_{ij}^2$. This shows that (14) follows an F-distribution on 1 and $(n-1)$ degrees of freedom.[2]

Hamaker [1955] has given an example showing that the sum of squares due to $\hat{c}$ is more effective than the sum of squares due to ${}_1\hat{c}$ in accounting for non-additivity. But Tukey's test has the advantage of being computationally simpler and of resulting in a test statistic that is known to follow the F-distribution. It should be noted that for this test we have to assume a model that allows for no interaction effect other than can be accounted for by a term $c[\alpha_i\beta_j]_{ij}$. This is better

[2]A more complete proof of this is given on p. 132 of H. Scheffé's book "The Analysis of Variance" [Wiley, 1959]. Scheffé also gives another motivation for Tukey's test; this came to the author's notice after the present paper was written.

than assuming a model that allows for no interaction at all, but not as good as we might wish. If we have more than one observation per subclass, we can assume the more reasonable model

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk} ,$$

$$\epsilon_{ijk} \text{ independent, } N(0, \sigma^2), \quad \text{and} \quad (\alpha\beta)_{ij} = c[\alpha_i\beta_j]_{ij} + \overline{(\alpha\beta)}_{ij}$$

and test the null hypothesis that $c = 0$, i.e. that

$$E(y_{ijk}) = \mu + \alpha_i + \beta_j + \overline{(\alpha\beta)}_{ij} ,$$

$\overline{(\alpha\beta)}_{ij}$ being that part of the interaction effect that cannot be accounted for by a term $c[\alpha_i\beta_j]_{ij}$. Then we can use the test statistic (assuming that there are n degrees of freedom associated with the within subclass sum of squares)

$$F = \frac{R(\text{II}') - R(\text{I})}{(\text{within subclass s.s.})/n} ;$$

this follows the F-distribution on 1 and n degrees of freedom under the null hypothesis, and, if condition (s) hold, provides an approximate test of whether the interaction sum of squares can be significantly reduced by a transformation of the data (using a monotonic function whose inverse is a quadratic polynomial).

The above can be easily generalized. For example, if we have a three-way classification, one observation per subclass, (6) becomes

$$E(y_{ijk}) = \mu + \alpha_i + \beta_j + \gamma_k + c(\alpha_i\beta_j + \alpha_i\gamma_k + \beta_j\gamma_k),$$

using an obvious notation. The sum of squares due to ${}_1\hat{c}$ is then

$$R(\text{II}') - R(\text{I}) = \frac{[\sum_{ijk} (\tilde{\alpha}_i\tilde{\beta}_j + \tilde{\alpha}_i\tilde{\gamma}_k + \tilde{\beta}_j\tilde{\gamma}_k)\bar{\epsilon}_{ijk}]^2}{\sum_{ijk} (\tilde{\alpha}_i\tilde{\beta}_j + \tilde{\alpha}_i\tilde{\gamma}_k + \tilde{\beta}_j\tilde{\gamma}_k)^2} ,$$

provided $\sum_i \tilde{\alpha}_i = \sum_j \tilde{\beta}_j = \sum_k \tilde{\gamma}_k = 0$. If there is more than one observation per subclass, the summations are extended over all observations in each subclass.

When there are unequal subclass numbers, it is often inconvenient computationally to use the restriction that the class effects should sum to zero (this is especially the case when a large body of data is being analyzed with the aid of an electronic computer). Analogous to equation (11) we can find, for the two-way classification, the restriction-free formula

$${}_1\hat{c} = \frac{\sum (y_{ijk}[\tilde{\alpha}_i\tilde{\beta}_j]_{ij} - \tilde{\mu}[\tilde{\alpha}_i\tilde{\beta}_j]_{ij} - \tilde{\alpha}_i[\tilde{\alpha}_i\tilde{\beta}_j]_{ij} - \tilde{\beta}_j[\tilde{\alpha}_i\tilde{\beta}_j]_{ij})}{\sum [\tilde{\alpha}_i\tilde{\beta}_j]^2_{ij}}$$

$$= \sum (y_{ijk}\tilde{\alpha}_i\tilde{\beta}_j - \tilde{\mu}\tilde{\alpha}_i\tilde{\beta}_j - \tilde{\alpha}_i^2\tilde{\beta}_j - \tilde{\alpha}_i\tilde{\beta}_j^2) / \sum [\tilde{\alpha}_i\tilde{\beta}_j]^2_{ij} ,$$

the summations being over all observations. Substituting this value into the equation obtained from (12) on replacing $\tilde{\alpha}_i\tilde{\beta}_j$ by $[\tilde{\alpha}_i\tilde{\beta}_j]_{ij}$ (and allowing for more than one observation per subclass), we obtain

$$R(\text{II}'') - R(\text{I}) = \frac{[\sum (y_{ijk}\tilde{\alpha}_i\tilde{\beta}_j - \tilde{\mu}\tilde{\alpha}_i\tilde{\beta}_j - \tilde{\alpha}_i^2\tilde{\beta}_j - \tilde{\alpha}_i\tilde{\beta}_j^2))]^2}{\sum [\tilde{\alpha}_i\tilde{\beta}_j]_{ij}^2},$$

the summations again being over all observations. The denominator in this expression is that part of $\sum (\tilde{\alpha}_i\tilde{\beta}_j)^2$ that is due to an interaction effect; it is conveniently obtained in a way analogous to that by which the interaction sum of squares of the y_{ijk} is obtained, but with the variable $\tilde{\alpha}_i\tilde{\beta}_j$ replacing y_{ijk} for all k.

FINDING AN APPROPRIATE TRANSFORMATION FOR ADDITIVITY

Consider, as a function of μ_a, $E(y_a) = \mu + \mu_a + c\mu_a^2$. If $c > 0$, this is a convex function, while, if $c < 0$, this is a concave function. Also, the inverse function given by

$$\mu_a = \{-1 + \sqrt{1 - 4c[(\mu - E(y_a)]}\}/2c, \tag{15}$$

is conversely (where it is real-valued) concave if $c > 0$ and convex if $c < 0$.

Note that we can write the inverse function (15)

$$L(\mu_a) = [(E(y_a) + (1/4c) - \mu]^{1/2},$$

where $L(\mu_a)$ is a linear function of μ_a. This would suggest that, if we reject the hypothesis that $c = 0$, a transformation that might be expected to bring our data closer to additivity is the transformation $x = (y + k)^{1/2}$ where y is the observed value and k is estimated by $1/4\hat{c} - \hat{\mu}$. However, in view of the fact that the inverse of model (II') (exemplified by equation (6) for the 2-way classification) is slightly different from (15), and that in any case such a model is unlikely to offer the best fit to the data, attention should not be restricted to this transformation alone. Tukey [1949a] suggests that we empirically seek out the best transformation in the class $x = (y + k)^p$, choosing $p < 1$ if ${}_1\hat{c} > 0$, and $p > 1$ [or trying $x = \log (y + k)$] if ${}_1\hat{c} < 0$. This advice regarding the choice of p follows from the fact that just as (15) is concave or convex according as $c > 0$ or $c < 0$, so $x = (y + k)^p$ is concave or convex according as $p < 1$ or $p > 1$. It is further suggested here that values of k approximating $(4\,|\,{}_1\hat{c}\,|)^{-2p} - \hat{\mu}$ be tried first, since this is a solution to

$$[E(y_a) + (1/4_1\hat{c}) - \hat{\mu}]^{1/2} = [E(y_a) + k]^p,$$

when $E(y_a)$ is replaced by $\hat{\mu}$. However, it has been found empirically,

from some artificially constructed examples, that this method of choosing k should only be used as a rough guide.

Now suppose we have two different transformations each of which we think may help in obtaining additivity. (For example, we might have two choices of p and k for the above general class of transformations; or we might have another entirely different transformation suggested by the nature of the data.) Then we might reasonably try and find the best linear combination of these two, i.e. a transformation of the form

$$x = d_1 f_1(y) + d_2 f_2(y),$$

where f_1 and f_2 are the two given functions and d_1 and d_2 are unknown coefficients. Tukey [1949b] has given a procedure for choosing d_1/d_2. His method can be extended to combine linearly more than two functions of the data, and the general method of obtaining suitable coefficients d_i will be given here.

Let $x = d_1 f_1(y) + d_2 f_2(y) + \cdots + d_p f_p(y)$. We will choose the p coefficients d_i given by the row vector $\mathbf{d}' = (d_1, d_2, \cdots, d_p)$ to maximize a certain ratio of sums of squares of the x's. Thus in a simple two-way classification Tukey considers maximizing

$$\frac{\text{row sum of squares of the } x\text{'s} + \text{column sum of squares of the } x\text{'s}}{\text{residual sum of squares of the } x\text{'s}}.$$

The method can be used to maximize any such ratio; and, provided we have a within-subclass sum of squares available, a general criterion would be to maximize

$$\frac{\text{within-subclass sum of squares of the } x\text{'s}}{\text{all interaction sums of squares of the } x\text{'s}},$$

since this would minimize any interaction variance component.

Let $s_1(f_i f_i)$ and $s_2(f_i f_i)$ be the numerator and denominator sum of squares respectively of the transformed data $f_i(y)$ in the ratio it is desired to maximize. Thus if we use the general criterion suggested above, $s_1(f_i f_i)$ is the within subclass sum of squares computed from the transformed data $f_i(y)$, and $s_2(f_i f_i)$ comprises all interaction sums of squares as computed after using the same transformation. Let $s_1(f_i f_j)$ and $s_2(f_i f_j)$ be respectively the corresponding numerator and denominator "sum of cross-products" of $f_i(y)$ and $f_j(y)$. For example, if $s_1(f_i f_i)$ were a row sum of squares, $s_1(f_i f_j)$ would be obtained as the sum of cross-products of row sums, divided by the usual divisor and less a corresponding correction term for the mean. If $s_1(f_i f_i)$ is the within subclass sum of squares, then $s_1(f_i f_j)$ is obtained exactly analo-

gously, except that $f_r(y)$ and $f_i(y)$ are used to obtain cross-products instead of just $f_i(y)$ to obtain squares.

Let $\mathbf{S}_1$ and $\mathbf{S}_2$ be $p \times p$ matrices whose elements in the i-th row and j-th column are $s_1(f_if_j)$ and $s_2(f_if_j)$ respectively.
Then

$$s_1(xx) = \sum_{i=1}^{p} d_i^2 s_1(f_if_i) + 2 \sum_{i=1}^{p} \sum_{j>i} d_i\, d_j s_1(f_if_j) = \mathbf{d}'\mathbf{S}_1\,\mathbf{d},$$

where $s_1(xx)$ is the numerator sum of squares of the x's and $\mathbf{d}$ is a column vector, the transpose of $\mathbf{d}'$. Similarly

$$s_2(xx) = \mathbf{d}'\mathbf{S}_2\,\mathbf{d}.$$

We thus wish to maximize $\mathbf{d}'\mathbf{S}_1\mathbf{d}/\mathbf{d}'\mathbf{S}_2\mathbf{d}$. Differentiating this with respect to $\mathbf{d}$ and equating to zero we obtain p equations that we may write in vector notation:

$$2\frac{\mathbf{S}_1\,\mathbf{d}}{\mathbf{d}'\mathbf{S}_2\,\mathbf{d}} - 2\frac{\mathbf{d}'\mathbf{S}_1\,\mathbf{d}}{(\mathbf{d}'\mathbf{S}_2\,\mathbf{d})^2}\mathbf{S}_2\,\mathbf{d} = \mathbf{0},$$

where $\mathbf{0}$ is a column vector of p zeros; i.e.

$$\mathbf{S}_1\,\mathbf{d} - \lambda\mathbf{S}_2\,\mathbf{d} = \mathbf{0} \tag{16}$$

where

$$\lambda = \frac{\mathbf{d}'\mathbf{S}_1\,\mathbf{d}}{\mathbf{d}'\mathbf{S}_2\,\mathbf{d}}.$$

Thus the largest value of λ is the value of the largest characteristic root of $(\mathbf{S}_1\mathbf{S}_2^{-1})$, and using this value of λ we choose $\mathbf{d}$ to satisfy (16). The individual d_i are not uniquely determined by (16), but all ratios d_i/d_j are. If we denote the element in the i-th row and j-th column of $(\mathbf{S}_1 - \lambda\mathbf{S}_2)$ by s_{ij}, we can easily find, for example when $p = 2$:

$$d_1/d_2 = -s_{12}/s_{11}\ ;$$

and when $p = 3$:

$$d_1/d_2 = (s_{13}s_{22} - s_{12}s_{23})/(s_{11}s_{23} - s_{12}s_{13}),$$

$$d_1/d_3 = (s_{12}s_{33} - s_{13}s_{23})/(s_{11}s_{23} - s_{12}s_{13}).$$

If we can find a $\mathbf{d}$ which makes all interaction sums of squares of the x's non-significant, we might assume the $E(x)$ are additive and so estimate our class effects on this scale. Unfortunately, however, it is not clear what hypotheses can be tested: any F-statistics we may obtain are influenced by our choice of $\mathbf{d}$. This procedure is therefore

best restricted to the case where two independent sets of similar data are available; $\mathbf{d}$ can be estimated from one set of data, and using this $\mathbf{d}$ the other set of data can be transformed and analyzed. Then, if we assume normality of the transformed data, tests of significance will be valid; though only one set of the data should be used in making them.

The computational procedure in the case of unequal subclass numbers will not be discussed in any detail here. However it should be noted that a general solution has already been given, in another context, by Roy [1957]. If we equate Roy's $\mathbf{x}'_r$ to the vector of transformed variables associated with the r-th observation y_r , i.e. to $[f_1(y_r), f_2(y_r), \cdots, f_p(y_r)]$, then $\mathbf{S}_1$ and $\mathbf{S}_2$ in this paper can be equated to Roy's $\mathbf{S}$ and $\mathbf{S}^*$ respectively.

REFERENCES

Hamaker, H. C. [1955]. Experimental design in industry. *Biometrics 11*, 257.

Roy, S. N. [1957]. *Some aspects of multivariate analysis*, 82–83. John Wiley and Sons, New York.

Tukey, J. W. [1949a]. One degree of freedom for non-additivity. *Biometrics 5*, 232.

Tukey, J. W. [1949b]. Dyadic anova, an analysis of variance for vectors. *Human Biology 21*, 65.

Tukey, J. W. [1955]. Answer to Query 113. *Biometrics 11*, 111.

Ward, A. C. and I. D. Dick [1952]. Non-additivity in randomized block designs and balanced incomplete block designs. *New Zealand Jour. Sci. and Tech. 33*, 430.

A NOTE ON SOME GROWTH PATTERNS IN A SIMPLE THEORETICAL ORGANISM

J. A. NELDER

National Vegetable Research Station, Wellesbourne, Warwick, England

INTRODUCTION

Hinshelwood and his co-workers [1946, 1955] have considered simple enzyme systems whose behaviour might mimic the growth of bacterial cultures under certain conditions. It is well known that the bacterial growth cycle of cell number on a given volume of medium divides itself into three main phases (Figure I). The first phase (or lag phase) represents a 'settling-in period' when growth may be slight or irregular; in the second phase, growth is logarithmic (that is log number is linearly related to time); finally in the third phase the relative growth rate falls from its constant value in the second phase to zero, often very rapidly, and the number in the colony becomes more or less steady. The lag time on transfer to a new culture may be generally defined by extending the linear portion of the logarithmic phase to the time axis (OT in Figure I). Considerable interest attaches to the form of the lag phase following (a) the transfer of cells from an ageing culture (i.e., one in the third phase) to a new medium and (b) the addition

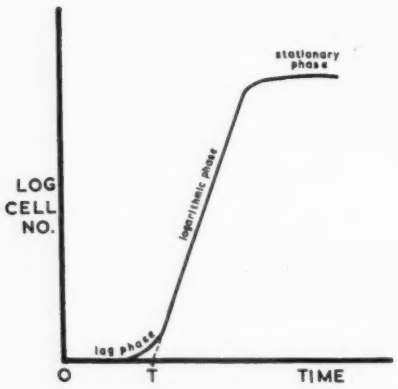

FIGURE I

IDEALIZED FORM OF THE RELATION BETWEEN LOG. NUMBER OF CELLS AND TIME IN A BACTERIAL CULTURE

of a drug to the new medium. Hinshelwood has considered [1946, chap. IV *et seq.*] forms of lag produced in a very simple theoretical organism consisting of two linked enzymes, I and II, both of which are autocatalytic; thus enzyme I uses substrate outside the organism to produce more of itself and also an intermediate C, while II uses C to produce more of itself. Cell division is assumed to be geared to the amount of II. One of his equations of growth, however, appears to need amendment, and it is fortunate that with this amendment the equations have an explicit solution so that it becomes possible to replace his approximate treatment of the lag phase by an exact one in a number of different situations.

THE EQUATIONS OF GROWTH

We write X_1 and X_2 for the total amount of enzymes I and II in the system at any time, n for the total number of cells, Y for the total amount of intermediate, and $y = Y/n$ for the concentration of intermediate on a per cell basis. The equations for the growth of the enzymes are then given by

$$\frac{dX_1}{dt} = k_1 X_1 , \tag{1}$$

$$\frac{dX_2}{dt} = k_2 X_2 y, \tag{2}$$

while the growth of the intermediate is given by

$$\frac{dY}{dt} = AX_1 - BX_2 y - CY, \tag{3}$$

where AX_1 is the rate of production by enzyme I, $-BX_2y$ the rate of consumption by enzyme II, and CY the loss by diffusion and irrelevant chemical processes, which must be proportional to the concentration and the number of cells n. (The notation is the same as that in Dean and Hinshelwood [1955] except that we have used y for c and have introduced $Y = nc$ in their notation). Equation (3) differs essentially from the corresponding one given by Dean and Hinshelwood in having $dY/dt = d(nc)/dt$ instead of $n\ dc/dt$. The difference arises because their equation does not take into account the fall in concentration of the intermediate due to the expansion of the system. This fall will be given by $-(y/n)(dn/dt)$ in our notation, so that in terms of concentration the equation for y is given by

$$\frac{dy}{dt} = A\frac{X_1}{n} - B\frac{X_2}{n}\,y - Cy - \frac{y}{n}\frac{dn}{dt}$$

whence

$$n\frac{dy}{dt} + y\frac{dn}{dt} = \frac{d}{dt}(ny) = \frac{dY}{dt} = AX_1 - BX_2y - CY,$$

thus reducing to (3), as it must.

Hinshelwood postulates that the amount of enzyme II controls cell division which thus acts so as to keep X_2/n, the concentration of II, roughly constant. In what follows, we follow him in ignoring the changes in X_2/n and the diffusion coefficient C due to the discontinuous nature of the division process, and assume that the system can be taken as expanding continuously in such a way that X_2/n is constant, also that the area/volume ratio is constant so that C is unrelated to n. We therefore put $X_2 = n\theta$ where θ is a constant (and equal to β^{-1} in Dean and Hinshelwood's notation), whereupon equations (1), (2) and (3) become the linear set

$$\frac{dX_1}{dt} = k_1X_1\,, \qquad \frac{dX_2}{dt} = k_2\theta Y, \qquad \frac{dY}{dt} = AX_1 - Y(B\theta + C). \tag{4}$$

The general solution of (4) may be found by the usual methods and gives

$$\begin{aligned} X_1 &= Q_1k_1(\alpha + k_1)e^{k_1t}, \\ X_2 &= Q_1k_2\theta A e^{k_1t} + Q_2 - Q_3k_2\theta e^{-\alpha t}, \\ Y &= Q_1k_1Ae^{k_1t} + Q_3\alpha e^{-\alpha t}, \end{aligned} \tag{5}$$

where Q_1, Q_2, and Q_3 are arbitrary constants and $\alpha = B\theta + C$. As t becomes large the organism tends to a steady state where X_1, X_2, and Y are all growing logarithmically in a constant ratio given by

$$X_1 : X_2 : Y = k_1(\alpha + k_1) : k_2\theta A : k_1A \tag{6}$$

When the organism is set going from a state different from the steady state given by (6), a lag phase will appear before the organism settles down into the logarithmic growth pattern. Two types of lag will be considered: (a) those produced by a decay in X_1 or Y (these possibilities are envisaged by Hinshelwood as being consequences of ageing in a culture) and (b) those produced by a drug which changes the value of one of the parameters in (4).

GROWTH FOLLOWING THE DECAY OF ENZYME AND INTERMEDIATE

(i) *The decay of* X_1

Suppose that at $t = 0$, $X_1 : X_2 : Y = pk_1(\alpha + k_1) : k_2\theta A : k_1A$, i.e. that relative to X_2 and Y, X_1 has declined to a fraction p of the steady

state ratio given by (6). Solving for Q_1 , Q_2 and Q_3 in (5) we find

$$n = X_2/\theta = \text{constant} \times \left\{\left(p/k_1 e\right)^{k_1 t} + (1-p)\left(\frac{1}{k_1} + \frac{1}{\alpha}\right) - \frac{1-p}{\alpha} e^{-\alpha t}\right\}. \quad (7)$$

Without loss of generality we may define the unit of n so that $n = 1$ at $t = 0$ and the unit of t so that $k_1 = 1$. Then (7) becomes

$$n = pe^t + (1-p) + \frac{1-p}{\alpha}(1 - e^{-\alpha t}). \quad (8)$$

As t becomes large

$$n \sim pe^t, \qquad \ln n \sim \ln p + t.$$

Hence the lag time T as defined above is obtained by putting $\ln n = 0$, whence $T = -\ln p$. Now α is a positive quantity in (8) and $(1 - e^{-\alpha t})/\alpha$ is a non-increasing function of α for any positive t; as $\alpha \to 0$, $(1 - e^{-\alpha t})/\alpha \to t$ while as $\alpha \to \infty$, $(1 - e^{-\alpha t})/\alpha \to 0$ so that extreme curves of the family (8) for fixed p are given by

$$\text{and} \qquad \begin{aligned} n &= pe^t + (1-p)(1+t), \\ n &= pe^t + (1-p). \end{aligned} \quad (10)$$

(ii) *The decay of Y*

If the amount of intermediate Y decays to a fraction p in the steady-state ratio (6) while the enzyme amounts remain fixed, then we obtain, corresponding to (8), the equation

$$n = e^t - [(1-p)(1 - e^{-\alpha t})/\alpha].$$

The lag time is thus zero and the curves lie between

$$n = e^t - (1-p)t, \quad (\alpha = 0), \quad \text{and} \quad n = e^t, \quad (\alpha = \infty).$$

GROWTH FOLLOWING CHANGES IN PARAMETER VALUES

One possible explanation of drug action is that it changes values of the parameters in the growth equations; thus an intermediate may be inactivated by the drug and be rendered more or less unavailable to the next enzyme in the chain, or active sites on an enzyme may be blocked, and so on. In the simple organism considered here growth depends on 5 parameters, k_1 , k_2 , θ, A, and $\alpha = B\theta + C$. We shall not consider variations in θ or k_1 ; if θ varied, then the drug would change the mean cell size and so alter C as well, while if k_1 is permanently changed, the final steady-state growth rate will be changed. We thus

consider effects on k_2 , A, and α which do not change the mean cell size or affect the steady-state growth rate after adaptation. In addition the drug will be assumed to be present in excess so that the effect on the parameters can be taken as fixed and independent of time. The results are summarised below:

(iii) *Fall in A*

The drug affects the production of intermediate by enzyme I. If A is reduced to a fraction p of its former value and cells are in their logarithmic phase before treatment then, following the same conventions as before

$$n = pe^t + (1 - p) + [(1 - p)(1 - e^{-\alpha t})/\alpha]$$

so that the effect is identical to that of a corresponding fall in the amount of X_1 (equation 8).

(iv) *Fall in k_2*

Here the drug affects the autocatalysis of enzyme II, and

$$n = pe^t + (1 - p),$$

which is an extreme case, as $\alpha \to \infty$, of equation (8). The lag in both these cases is given by $T = -\ln p$.

(v) *Rise in α*

In this case the drug is assumed to mop up the intermediate as it is produced, thus increasing C and hence α. The quantity α can also be increased by increasing B, the rate of use of the intermediate by enzyme II, but this seems a less likely effect of drug action. The cell number is given by

$$n = \frac{1}{1 + p\alpha}\left[(1 + \alpha)e^t - \alpha(1 - p) - \frac{(1 - p)}{p}(1 - e^{-p\alpha t})\right] \qquad (11)$$

where $p > 1$, since α is assumed to have increased. The lag is given by $T = \ln [(1 + p\alpha)/(1 + \alpha)]$. The extreme members of (11) are given

$$n = e^t \quad \text{for} \quad \alpha = 0, \qquad (12)$$

and

$$n = \pi e^t + (1 - \pi) \quad \text{for} \quad \alpha = \infty, \qquad \pi = 1/p. \qquad (13)$$

Equation (12) of course gives logarithmic growth with no lag, while (13) has been met previously as one of equations (10).

GROWTH FOLLOWING DISTURBANCE OF TWO QUANTITIES OR PARAMETERS

In this section we consider only the growth following the disappearance of intermediate accompanied by changes in X_1 or one of the parameters A, k_2, or α. Three of these cases produce the same equation for n, namely those having a fall in X_1, A, or k_2 to a fraction p of their steady state value. All these give

$$n = pe^t + (1 - p) - [p(1 - e^{-\alpha t})/\alpha],$$

with lag $T = -\ln p$, and extreme curves

$$n = pe^t + (1 - p) - pt \quad \text{for} \quad \alpha = 0,$$
$$n = pe^t + (1 - p) \qquad \text{for} \quad \alpha = \infty.$$

The effect of α is not large here; the curves having $\alpha = 0$ have a slightly more sudden rise to the logarithmic asymptote.

For a rise in α to $p\alpha$ accompanied by an initial value of $Y = 0$ we get

$$n = \frac{1}{1 + p\alpha}\left[(\alpha + 1)e^t + (p - 1)\alpha - \frac{\alpha + 1}{p\alpha}(1 - e^{-p\alpha t})\right].$$

The lag is $T = \ln [(1 + p\alpha)/(1 + \alpha)]$ and the extreme curves are given by

$$n = e^t - t \quad \text{for} \quad \alpha = 0,$$
$$n = \pi e^t + (1 - \pi), \qquad \pi = 1/p \quad \text{for} \quad \alpha = \infty.$$

EXAMPLES OF GROWTH CURVES

Cases (i) and (iii) give the same family of curves. Examples of these are shown in Figure II (a), (b), and (c) for the extreme values of α (0 and ∞), and for $\alpha = 1$, the curves being given for lags equal to 2, 4, 8, and 16 generation times. (On our standard scale with $k_1 = 1$, the generation time is $\ln 2 = 0.693$ units.) Case (iv) is covered by Figure II(b), while case (v) has as extremes the simple zero-lag line $\ln n = t$ for $\alpha = 0$ and the curves of Figure II(b) for $\alpha = \infty$. For $\alpha = 1$, curves very similar to those of Figure II(b) are obtained. In Figure II(d) an example is given for case (ii) with the extreme values $\alpha = 0$, $p = 0$; for $p < 1$, these curves are sigmoid and asymptotic to the no-lag line $\ln n = t$ from below, the slope at the origin being p for all finite α.

For growth following complete loss of Y accompanied by a change in X_1, A, k_2 or α, the curves are similar to those of Figure II(b) for all positive α, and examples are not given.

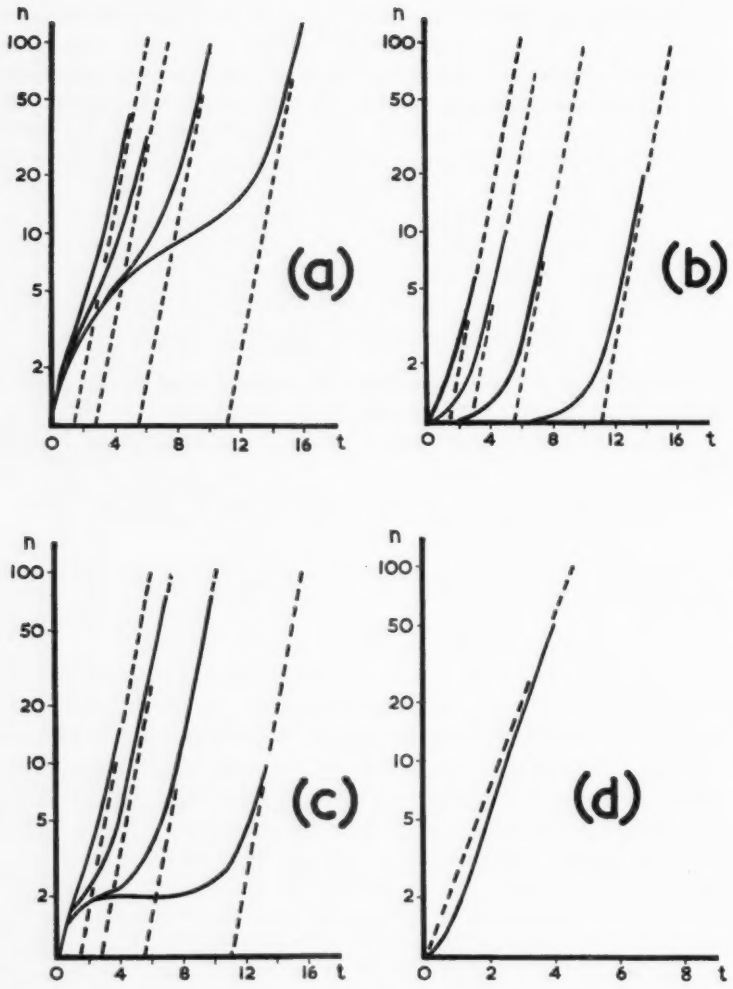

FIGURE II

EXAMPLES OF GROWTH CURVES AFTER DISTURBANCE FROM THE STEADY STATE

(a) Curves of the family $n = pe^t + (1 - p)(1 + t)$
(b) Curves of the family $n = pe^t + (1 - p)$
(c) Curves of the family $n = pe^t + (1 - p)(2 - e^{-t})$
(d) The curve $n = e^t - t$

Notes: the ordinate is ln n throughout. In (a), (b), and (c) p takes the values 2^{-2}, 2^{-4}, 2^{-8}, and 2^{-16}.

DISCUSSION

Though the "organism" discussed in this note is a very simple one, the curves show that it is capable of a variety of different growth patterns after disturbance from the steady state. The curves exemplified in Figure II(c) are somewhat similar to one of those shown in Figure 10 of Hinshelwood's book (Hinshelwood [1946]), and described by him as an "irregular growth curve". However we have not succeeded in producing a curve such as that shown in Figure 11 of the same work where growth is very slight for about 4 generation times and the logarithmic phase is then entered very rapidly indeed. When the lag is as short as 4 generation times, none of the curves produced by the model has a near-horizontal early part. It is only when the lag reaches about 10 generation times that we find in some cases a substantial period at the beginning when growth is very slight.

One way of producing a curve of the type of Hinshelwood's Figure 11 might be to postulate a drug which combined with the intermediate and prevented its use by enzyme II, the drug itself becoming inactivated by the combination. In this situation the amount of intermediate available to enzyme II would be kept very low until all the drug had been immobilized after which it would rise to the stable state level. The general question of the behaviour of the organism 'growing away' from a drug after immobilizing it is complex and does not seem amenable to exact solution. Similarly the growth equations for Hinshelwood's other theoretical organisms, such as the one having two cyclically linked enzymes (Hinshelwood [1946], p. 81 *et seq.*), or alternative pathways (*ibid.* p. 150 *et seq.*) appear to have no explicit general solution, though of course numerical solutions for given values of the constants could easily be obtained.

A striking feature of the model is the occurrence of similar or identical growth equations of n following different changes of the steady state. Since these changes have quite distinct biological interpretations, it is clear that it will rarely be possible to deduce from the shape of a growth curve the sort of mechanism likely to have produced it.

SUMMARY

The growth equations for a simple organism proposed by Hinshelwood, and consisting of two linked autocatalytic enzymes, are shown to produce a linear system under certain conditions.

Solution of these equations gives the form of the lag in growth following (i) decay of one enzyme and/or the intermediate and (ii) changes in the values of the parameters such as might be brought about by drug action.

Examples of the types of curves produced are given, and their resemblance to actual bacterial growth curves discussed. The production of identical curves by several distinct mechanisms in the model is pointed out.

REFERENCES

Hinshelwood, Sir Cyril [1946]. *The Chemical Kinetics of the Bacterial Cell.* Oxford University Press, Oxford, England.

Dean, A. C. R., and Hinshelwood, Sir Cyril [1955]. Reaction Patterns of a Coliform Organism. *Progressive Biophysics & Biophys. Chem.* 5, 1–40.

THE PARTIAL DIALLEL CROSS[1]

O. KEMPTHORNE
Iowa State University, Ames, Iowa, U. S. A.
AND
R. N. CURNOW[2]
Agricultural Research Council Unit of Statistics, University of Aberdeen, Aberdeen, Scotland.

1. *Introduction*

The diallel cross, which is composed of all possible single crosses among a group of inbred lines, is now a common plan of investigation in both plant and animal breeding. Its modern use starts apparently with the development of the concepts of general and specific combining ability by Sprague and Tatum [1942]. The diallel cross is used to estimate the genetic components of the variation among the yields of the crosses (see, for instance, Hayman [1954a, b]; Jinks and Hayman [1953]; Griffing [1956a, b] and Kempthorne [1956, 1957]). It is also used to estimate the actual yielding capacities of the crosses. This information may be employed, for example, to select the best four inbred lines from which to develop a four-way cross. We shall discuss only the relatively simple situation in which there are no maternal effects, i.e. reciprocal crosses are identical and so need not be made, and in which there is no interest in the performance of the inbred lines themselves. The important questions concerned with replication over years and locations will not be considered. The methods used and conclusions reached in this paper may be capable of extension to more complicated situations.

With no reciprocal crosses or crosses resulting from selfing or crossing within the same inbred line, there are ${}^nC_2 = n(n-1)/2$ possible single crosses among n inbred lines. This number of possible crosses increases rapidly with n. When $n = 6$ it is only 15, when $n = 20$ it is 190, and when $n = 50$ it is 1,225. With facilities available for testing only a limited number of crosses, a diallel cross may only be possible between a relatively small number of inbred lines. If only a small number of inbred lines are tested, the estimate of the variance of the

[1]Journal Paper No. J-3909 of the Iowa Agricultural and Home Economics Experiment Station, Ames, Iowa. Project 890.
[2]The work on this paper was completed while the latter author was visiting Iowa State University on a Harkness Fellowship of the Commonwealth Fund.

general combining abilities among the whole population of potentially available inbred lines will be subject to a large sampling error and many potentially high yielding inbred lines left completely untested. The question arises, therefore, of whether a design involving more inbred lines but only a sample of all possible crosses between them may not be preferable. A particular method of sampling the diallel cross will be considered and its efficiency in estimating the genetic variance components, the differences between the yields of the various crosses and the general combining abilities of the lines discussed.

2. *The Partial Diallel Cross*

The best method to be used in deciding how to sample the diallel cross would be to specify, in general terms, the probabilities of sampling each particular cross and each particular pair of crosses and then choose these probabilities so as best to satisfy the aims of the experiment. We have not been able to do this. We shall consider only the particular method of sampling the crosses suggested by G. W. Brown in about 1948. The method certainly achieves some balance and whether a much better one exists is rather doubtful. It has already been used twice at Iowa State University (Jensen, [1959]; Sprague, unpublished.).

Assume that the breeder can handle a total of $ns/2$ crosses where n is the number of inbred lines and s is a whole number greater than or equal to 2. Clearly, n and s cannot both be odd. The n inbred lines are numbered at random from 1 to n and the following crosses sampled:—

$$\begin{aligned}
&\text{line } 1 \times \text{lines } k+1, k+2, \cdots, k+s \\
&\text{line } 2 \times \text{lines } k+2, k+3, \cdots, k+1+s \\
&\quad\cdots \qquad\qquad \cdots \qquad\qquad \cdots \\
&\text{line } i \times \text{lines } k+i, k+i+1, \cdots, k+i-1+s \\
&\quad\cdots \qquad\qquad \cdots \qquad\qquad \cdots \\
&\text{line } n \times \text{lines } k+n, k+n+1, \cdots, k+n-1+s,
\end{aligned}$$

where $k = (n+1-s)/2$, and is a whole number, and all the numbers above n are to be reduced by multiples of n so as to be between 1 and n. For k to be a whole number, either n is odd and s even or n is even and s odd. Each line occurs in s crosses and the total number of crosses sampled is $ns/2$. $s = n - 1$ corresponds to the complete diallel cross. The reader may have noticed the analogy between this method of sampling the diallel cross and the experimental design for blocks of two plots in which the s blocks containing treatment number i also contain treatments numbered $k+i, k+i+1, \cdots, k+i-1+s$. These circulant designs are discussed by Kempthorne [1953]. To

every design for blocks of two plots in which each treatment occurs the same number of times, there corresponds a method of sampling the diallel cross in which each line is involved in the same number of crosses, and conversely. Treatments i and j occurring in the same block corresponds to sampling the cross $i \times j$. Any balanced incomplete block design corresponds to the complete diallel cross. Partial diallel crosses corresponding to other two plot per block designs have not been considered. Partial diallel crosses corresponding to other circulant designs, including some with n and s both even, (Zoellner and Kempthorne, [1954]), could be studied by methods very similar to those of this paper.

We shall assume that the yield from the cross $i \times j$ in replicate l can be written

$$y_{ijl} = \mu + r_l + g_i + g_j + s_{ij} + e_{ijl} ,$$

where μ is a general effect, r_l a replicate effect, g_i and g_j are the parental effects (sometimes called general combining abilities), s_{ij} is the effect of the non-additivity of the parental effects (sometimes called specific combining ability) and e_{ijl} is the plot error. We shall further assume that g_i , s_{ij} and e_{ijl} are independently normally distributed with zero means and variances σ_g^2 , σ_s^2 and σ_e^2 . The motivation for this model is as follows. The yield of cross ij in replicate l is assumed to consist of three parts combining additively: an effect of the cross, a replicate effect and an error due to plot deviation and also to segregation within the cross, if there is any. As regards the cross effect, different progeny of the same cross are full sibs, and progeny of two different crosses involving a common line are half sibs, and this is the only genetical structure. If the lines are a random sample from a large population, the cross effect can be represented by

$$g_i + g_j + s_{ij}$$

in which g_i , g_j and s_{ij} are assumed to be independent random variables with variances σ_g^2 , σ_g^2 and σ_s^2 respectively, where

$$\sigma_g^2 = \text{Cov (H.S.)}$$

and

$$\sigma_s^2 = \text{Cov (F.S.)} - 2\,\text{Cov (H.S.)},$$

H.S. denoting half-sibs and F.S. denoting full-sibs.

The diallel cross can be used with multi-flowered plants and, if the plants used are a random sample of the population of plants, the above covariances are covariances of relatives in that population. If further-

more that population has random mating structure, is not inbred and there is no linkage,

$$\text{Cov (H.S.)} = \tfrac{1}{4}\sigma_A^2 + \tfrac{1}{16}\sigma_{AA}^2 + \cdots$$

and

$$\text{Cov (F.S.)} = \tfrac{1}{2}\sigma_A^2 + \tfrac{1}{4}\sigma_D^2 + \tfrac{1}{4}\sigma_{AA}^2 + \cdots .$$

If the parents are individuals arising by inbreeding to degree F in such a population, then (Cockerham [1954], Kempthorne [1957]

$$\text{Cov (H.S.)} = \left(\frac{1+F}{4}\right)\sigma_A^2 + \left(\frac{1+F}{4}\right)^2\sigma_{AA}^2 + \cdots$$

and

$$\text{Cov (F.S.)} = \left(\frac{1+F}{2}\right)\sigma_A^2 + \left(\frac{1+F}{2}\right)^2\sigma_D^2 + \left(\frac{1+F}{2}\right)^2\sigma_{AA}^2 + \cdots .$$

If F equals unity, i.e. the parents are completely inbred, Cov (H.S.) is equal to the covariance of parent and offspring in the original random mating population and Cov (F.S.) is the genotypic variance in the original population. If the lines are only partially inbred, genetic interpretation is possible only if the degree of inbreeding is constant over the lines and is known.

For a general value of F, all we know about σ_g^2 and σ_s^2 is that the former, σ_g^2 = Cov (H.S.), involves only the variances due to additive effects and interactions of additive effects and that the latter, σ_s^2 = Cov (F.S.) − 2 Cov (H.S.), does not involve the variance due to additive effects, σ_A^2. If the lines are completely inbred,

$$\sigma_g^2 = \tfrac{1}{2}\sigma_A^2 + \tfrac{1}{4}\sigma_{AA}^2 + \cdots$$

and

$$\sigma_s^2 = \sigma_D^2 + \tfrac{1}{2}\sigma_{AA}^2 + \cdots .$$

The total genotypic variance is $2\sigma_g^2 + \sigma_s^2$ and, if epistacy can be ignored, the additive variance is $2\sigma_g^2$. The ratio of additive to total genotypic variance would then be

$$2\sigma_g^2/(2\sigma_g^2 + \sigma_s^2),$$

and the square root of twice the ratio of dominance to additive variance, which, if gene frequencies are equal to $\frac{1}{2}$, is interpretable as the average degree of dominance, is

$$\sqrt{\sigma_s^2/\sigma_g^2}\,.$$

If the lines are not inbred at all, i.e. $F = 0$,

$$\sigma_g^2 = \tfrac{1}{4}\sigma_A^2 + \tfrac{1}{16}\sigma_{AA}^2 + \cdots$$

and

$$\sigma_s^2 = \tfrac{1}{4}\sigma_D^2 + \tfrac{1}{8}\sigma_{AA}^2 + \cdots .$$

If epistacy can be ignored, the ratio of additive to total genotypic variance is

$$\sigma_g^2/(\sigma_g^2 + \sigma_s^2),$$

and the average degree of dominance, (see above)

$$\sqrt{2\sigma_s^2/\sigma_g^2} .$$

The plot error variance σ_e^2 will contain a component due to genetic variability within crosses. In fact it will contain the total genotypic variance less the covariance of full-sibs and this is zero only if F equals unity. The environmental variability in the plot error e_{ijl} consists of plot to plot variability plus environmental variability particular to each individual and, if there were competition, would also contain a component from this force.

The above model seems entirely appropriate for the consideration of the estimation of σ_g^2 and σ_s^2 , which are known as half the variance of general combining ability and the variance of specific combining ability respectively. Likewise it seems entirely appropriate for the estimation of the covariances of half-sibs and full-sibs, or estimation of related quantities like the average degree of dominance.

It is less appropriate for the consideration of the yields of possible multi-way crosses, such as two-way or four-way crosses, because the plant breeder will be in the position of having a specified set of lines and is interested in the potentialities within that set of lines rather than within a random set of lines that he might have obtained. But it appears to us that even in this case the plant breeder has no alternative to assuming an additive model in which the yield of cross $i \times j$ is made up of an effect due to line i, an effect due to line j, and an effect due to the non-additivity of the effects of lines i and j. Earlier an attempt was made to consider the situation in the framework that there was a finite population of N lines from which n were drawn at random and the partial diallel cross made among these n lines. Unfortunately, considerable mathematical difficulties were encountered and no neat result obtained. We shall therefore look at the problem of estimating all possible single crosses by the use of the simple additive model described above.

3. *Estimation of the Variance Components*

We shall estimate the general combining abilities $g_1, \cdots, g_n$ by least squares, i.e. by choosing $\hat{\mu}, \hat{g}_1, \cdots, \hat{g}_n$ to minimize

$$\sum_s (y_{ij} - \hat{\mu} - \hat{g}_i - \hat{g}_j)^2,$$

where $\sum_s$ denotes summation over the sampled crosses and y_{ij} the mean yield of cross $i \times j$. With the imposed constraint

$$\sum_{i=1}^{n} \hat{g}_i = 0,$$

this leads to the equation $\hat{\mu} = \bar{y}$, the average of the y_{ij}'s, and to the following equations for $\hat{g}_1, \cdots, \hat{g}_n$,

$$\sum_{j=1}^{n} a_{ij}\hat{g}_j = \sum_{r=0}^{s-1} \left[y_{i,i+k+r} - \frac{2G}{ns} \right] = Q_i, \qquad (i = 1, 2, \cdots, n), \quad (3.1)$$

where G is the grand total of cross averages and $a_{ii} = s$, all i, and $a_{ij} = a_{ji} = 1$ if cross $i \times j$ is sampled and $a_{ij} = a_{ji} = 0$ otherwise. The $n \times n$ matrix $\mathbf{A} = [a_{ij}]$ is a symmetric circulant matrix and therefore so is its inverse, $\mathbf{A}^{-1} = [a^{ij}]$. (The general form of a circulant matrix is given in the appendix at the end of this paper. A symmetric circulant matrix has the additional property that the element in the ith row and jth column is the same as the element in the jth row and ith column for all i and j.) Being a symmetric circulant, the elements a^{ij} of the matrix $\mathbf{A}^{-1}$ are functions only of $| i - j |$, the positive value of $(i - j)$. We shall therefore write $a^{ij} = a^{|i-j|}$. By multiplying together rows of $\mathbf{A}$ with columns of $\mathbf{A}^{-1}$,

$$\sum_{r=k}^{k+s-1} a^r = 1 - sa^0 \qquad (3.2)$$

and

$$\sum_{r=k}^{k+s-1} a^{t+r} = -sa^t, \qquad t = 1, \cdots, n-1. \qquad (3.3)$$

In (3.3), the $t + r$ index in the summation is to be reduced by multiples of n so as always to be between 0 and $n - 1$. By summing (3.3) from $t = 1$ to $t = n - 1$, and using (3.2), the sum of any row of $\mathbf{A}^{-1}$ is

$$\sum_{r=0}^{n-1} a^r = 1/2s. \qquad (3.4)$$

The total sum of squares of cross averages in the analysis of variance of the partial diallel cross can be decomposed as follows:

$$\sum_{s} (y_{ij} - \bar{y})^2 = \sum_{i=1}^{n} \hat{g}_i Q_i + \sum_{s} (y_{ij} - \bar{y} - \hat{g}_i - \hat{g}_j)^2,$$

where $\bar{y}$ is the mean of the cross averages.

Now from (3.1), and (3.4)

$$\sum_{i=1}^{n} \hat{g}_i Q_i = \sum_{i=1}^{n} \sum_{j=1}^{n} a^{ij} Q_i Q_j$$

$$= \sum_{i,j=1}^{n} \sum_{l,m=0}^{s-1} a^{ij} y_{i,i+k+l} y_{j,j+k+m} - 2G^2/ns. \qquad (3.5)$$

Fixing a particular sampled cross, $i \times (i + k + l)$, and considering all sampled crosses, $j \times (j + k + m)$, related to it, either as full or half sibs, and using (3.2), (3.3) and (3.4),

$$E\left(\sum_{j=1}^{n} \sum_{m=0}^{s-1} a^{ij} y_{i,i+k+l} y_{j,j+k+m}\right) = \frac{\mu^2}{2} + (a^0 + a^{k+l})$$

$$\cdot\left\{\frac{\sigma_e^2}{r} + \text{Cov (F.S.)} - 2\,\text{Cov (H.S.)}\right\} + \text{Cov (H.S.)}. \qquad (3.6)$$

Summing (3.6) over i and l, again using (3.2), (3.3) and (3.4), and substituting an expression for $E(G^2)$, (3.5) gives

$$E\left(\sum_{i=1}^{n} \hat{g}_i Q_i\right) = (n-1)\left\{\frac{\sigma_e^2}{r} + \text{Cov (F.S.)} - 2\,\text{Cov (H.S.)}\right\}$$

$$+ s(n-2)\,\text{Cov (H.S.)}$$

Table 1 shows the analysis of variance of the partial diallel cross, with the expected values of the mean squares written in terms of σ_g^2 and σ_s^2 instead of in terms of Cov (F.S.) and Cov (H.S.).

TABLE 1

ANALYSIS OF VARIANCE OF PARTIAL DIALLEL CROSS

Source	d.f.	Expected values of mean squares
Replicates	$r - 1$	
General Combining Ability	$n - 1$	$\sigma_e^2 + r\sigma_s^2 + \{rs(n-2)/(n-1)\}\,\sigma_g^2$
Specific Combining Ability	$n(s/2 - 1)$	$\sigma_e^2 + r\sigma_s^2$
Replicates × Crosses	$(r - 1)(ns/2 - 1)$	σ_e^2
Total	$rns/2 - 1$	

With a complete diallel cross there are always many more degrees of freedom for the s.c.a. (specific combining ability) mean square than for the g.c.a. (general combining ability) mean square. Unless σ_g^2 is small compared to $\sigma_e^2 + r\sigma_s^2$, this may be a serious disadvantage in the estimation of σ_g^2. The partial diallel cross does allow the $(ns/2 - 1)$ degrees of freedom for crosses to be split more evenly between the two mean squares. Clearly, for σ_s^2 to be estimable, s has to be greater than 2. With $s = 3$, there will be nearly twice as many degrees of freedom for the g.c.a. mean square as for the s.c.a. mean square. With $s = 4$, the degrees of freedom will be approximately equal. Table 2 shows the breakdown of the degrees of freedom for an experiment involving 2 replicates of 120 plots each when the crosses are (i) a complete diallel cross ($s = n - 1 = 15$), (ii) a partial diallel cross with $s = 3(n = 80)$ and (iii) a partial diallel cross with $s = 4(n = 60)$. The determination of the optimal numbers of replicates r, inbred lines n, and crosses per inbred line $s/2$ will involve the unknown values of σ_g^2/σ_e^2 and σ_s^2/σ_e^2 as well as the cost function of the experiment. Another difficulty in deciding on a value for s is that an exact specification of the aims of the experiment is necessary. One possible aim of the experiment is to estimate components of genetic variance such as the additive and dominance portions and the goodness of the design will depend on these and on the inbreeding coefficient of the lines. A compromise may have to be made between various aims. The proportion of the genetic variance that is additive (Fisher, [1918]) and the average degree of dominance (Comstock and Robinson, [1948, 1952]) are both, under severe assumptions (see Section 2), functions of σ_s^2/σ_g^2. Again, measures

TABLE 2

DEGREES OF FREEDOM FOR COMPLETE AND FOR PARTIAL DIALLEL CROSS USING 2 REPLICATES OF 120 PLOTS EACH

Source	Degrees of freedom		
	Complete diallel cross	Partial diallel cross	
		$s = 3$	$s = 4$
Replicates	1	1	1
G. c. a.	15	79	59
S. c. a.	104	40	60
Replicates × Crosses	119	119	119
Total	239	239	239

of heritability will be functions of σ_g^2 , σ_s^2 , and σ_e^2 . A design that minimizes the variance of the estimates of σ_g^2 and σ_s^2 does not necessarily minimize the variance of the estimates of these functions. For example, by increasing the degrees of freedom for the g.c.a. mean square at the expense of the degrees of freedom for the s.c.a. mean square, the variance of the estimate of σ_s^2 is increased and the covariance between the estimates of σ_g^2 and σ_s^2 made more negative. This may result in an increase in the variance of the estimate of σ_s^2/σ_g^2 , despite the decreases in the variance of the estimate of σ_g^2 .

4. *Comparison of Partial Diallel Cross with Other Designs for Estimating the Variance Components*

A comparison of the partial or complete diallel cross with other designs for estimating the variance components is complicated by all the factors mentioned at the end of the last section. The experiment analogous to what Comstock and Robinson [1952] called Experiment I would be to cross each of m randomly chosen inbred lines used as "sires" with m different sets of f randomly chosen inbred lines used as "dams". This type of experiment has been widely used in poultry to estimate covariance of full-sibs and half-sibs. No inbred line would be used both as a "sire" and a "dam". The analogy is not complete since the experimental material considered by Comstock and Robinson was produced from random matings among plants of the F_2 generation of a cross of two inbred lines, whereas we are considering the progeny resulting directly from crosses among a set of inbred lines. The analysis of variance of Experiment I is shown in Table 3. The variance of the estimates of both σ_g^2 and σ_s^2 with this design can be shown to be greater than the corresponding variances for the partial diallel cross with $n = m$ and $s = 2f$. The two designs involve the same number of

TABLE 3

ANALYSIS OF VARIANCE OF AN EXPERIMENT IN WHICH m INBRED LINES ARE EACH CROSSED TO A DIFFERENT SET OF f INBRED LINES

Source	d.f.	Expected values of mean squares
Replicates	$r - 1$	
"Sires"	$m - 1$	$\sigma_e^2 + r\sigma_s^2 + r(f + 1)\,\sigma_g^2$
"Dams in Sires"	$m(f - 1)$	$\sigma_e^2 + r\sigma_s^2 + r\sigma_g^2$
Replicates × Crosses	$(r - 1)(mf - 1)$	σ_e^2
Total	$rmf - 1$	

crosses but the partial diallel cross uses f fewer inbred lines. The covariance between the estimates of σ_g^2 and σ_s^2 is less negative with the diallel cross than with Experiment I. The experiment analogous to Experiment II of Comstock and Robinson [1952] would be to make all the crosses between m randomly chosen inbred lines used as "sires" and f randomly chosen inbred lines used as "dams". No inbred line would be used both as a "sire" and a "dam". The analysis of variance of this design is shown in Table 4. An estimation procedure that does not use both the sire and dam mean squares to estimate σ_g^2 can be shown to be inferior to the estimation procedure possible with a partial diallel cross with $n = m$ and $s = 2f$, i.e. a partial diallel cross involving the same number of crosses but f fewer inbred lines. Comstock and Robinson [1952] have suggested that, when possible, Experiment II should be designed with $m = f$. This is not necessarily optimal, but it does permit a simple pooling of the dam and sire mean squares. Unfortunately, it also results in many more degrees of freedom for the s.c.a. mean square than for the g.c.a. mean square. There will be $(m - 1)/2$ times as many degrees of freedom for the s.c.a. mean square as for the g.c.a. mean square. Unless σ_g^2 is small compared with $\sigma_e^2 + r\sigma_s^2$, this may be a serious disadvantage in the estimation of σ_g^2. Indeed one of the main general reasons for using the partial diallel cross is that it gives a reasonable number of degrees of freedom for the g.c.a. mean square.

A comparison of the partial diallel cross with Experiment III of Comstock and Robinson [1952] and other more complicated designs will not be attempted. Experiment III may be superior to the partial diallel cross. A complete comparison of the two designs would again involve the values of unknown parameters and would also raise questions

TABLE 4

ANALYSIS OF VARIANCE OF ALL CROSSES BETWEEN TWO DIFFERENT SETS OF m AND f INBRED LINES

Source	d.f.	Expected values of mean squares
Replicates	$r - 1$	
"Sires"	$m - 1$	$\sigma_e^2 + r\sigma_s^2 + rf\sigma_g^2$
"Dams"	$f - 1$	$\sigma_e^2 + r\sigma_s^2 + rm\sigma_g^2$
"Sires × Dams"	$(m - 1)(f - 1)$	$\sigma_e^2 + r\sigma_s^2$
Replicates × Crosses	$(r - 1)(mf - 1)$	σ_e^2
Total	$rmf - 1$	

regarding the sensitivity of the estimation procedures to departures from some of the genetic assumptions.

All the designs discussed so far have involved crosses between lines both of which were drawn at random from the population of inbred lines. One result of this is that all of the variation between crosses can be used to estimate either σ_g^2 or σ_s^2 . There are two serious disadvantages involved in the use of common testers that are not a random sample from the same population as the inbred lines. Firstly, the general combining ability of a line has to be defined as the average performance of that line when crossed to all members of a certain population. The average performance of an inbred line when crossed to a set of common testers will almost certainly not be the same as its average performance when crossed to all other members of the population of inbred lines. This could seriously invalidate the estimate of σ_g^2 . Secondly, there would seem to be little or no reason to assume that the variance of the specific combining abilities among crosses between inbred lines and common testers is equal to the variance of the specific combining abilities among crosses between the inbred lines themselves. These considerations strongly suggest that common testers should not be used to estimate the values of σ_g^2 and σ_s^2 for the population of inbred lines, unless the common testers are themselves a random sample from that same population.

5. *Comparing the Yielding Capacities of the Crosses*

We shall make the same assumptions and use the same notation as in the previous sections of this paper. We shall consider the partial diallel cross as a method of estimating the yielding capacities of all the possible single crosses among n inbred lines. Inbred lines are often developed and then crossed in an attempt to produce crosses with a high yielding capacity. The yielding capacity of all crosses in the diallel cross and all sampled crosses in the partial diallel cross can be estimated in each of two ways. Firstly, they can be estimated by their mean yields in the experiment. Secondly, specific combining ability can be ignored and the yielding capacity of the cross $i \times j$ estimated by $\hat{\mu} + \hat{g}_i + \hat{g}_j$.[3] Clearly, the latter method has to be used to estimate the yielding capacity of all the unsampled crosses in the partial diallel cross.

We shall designate the methods of estimation A and B, where A and B are defined as follows:

[3]The two estimates could possibly be given appropriate weights and combined into a single estimate. For the partial diallel cross these weights are rather involved and vary from cross to cross. In this paper, the two methods of estimation will be considered separately and then compared.

A. Unsampled crosses estimated by $\hat{\mu} + \hat{g}_i + \hat{g}_j$ but sampled crosses estimated by cross means y_{ij} .

B. Both unsampled and sampled crosses estimated by $\hat{\mu} + \hat{g}_i + \hat{g}_j$.

A more precise statement will be made later, but one may summarize briefly that *B* will be preferable to *A* only if $r\sigma_s^2/\sigma_e^2$ is small.

A third method *C* of estimating the yielding capacities of the crosses is to cross all n inbred lines with each of t common testers. As mentioned previously, general combining ability is not a property of the line alone but a property of the line relative to the population of lines to which it has been crossed. To speak of the general combining ability of a line without reference to that population is vague if not meaningless. If common testers are to be used to estimate the yielding capacities of crosses between the inbred lines then we have to make the assumption, and it is a big assumption, that the general combining ability of each inbred line relative to the common testors is the same as its general combining ability relative to all the other inbred lines. We hope that the importance of this assumption will not be lost in the discussion that follows. Without it there is no basis for comparing the diallel cross with common testers. In method *C*, we shall estimate the yielding capacity of cross $i \times j$ by $\hat{\mu} + \hat{g}_i + \hat{g}_j$ where $\hat{g}_i$ and $\hat{g}_j$ are the estimates of the general combining abilities of lines i and j with the common testers. Like *B*, *C* will be preferable to *A* only if $r\sigma_s^2/\sigma_e^2$ is small. Neither of the two designs alternative to the diallel cross considered in Section 4 can be used to estimate the differences between crosses without confounding some of these differences with other contrasts between the general combining abilities. They will therefore not be considered further.

In all three methods, *A*, *B* and *C*, the errors involved in comparing any two crosses will be composed of two parts. There will be errors arising from an inadequate estimation of specific combining abilities and from the specific combining abilities that enter into the estimates $\hat{g}_i$. There will also be errors arising from the plot effects. To measure the error in comparing any two crosses, we shall calculate the expected value of the square of this error. It will be composed of two parts, one involving σ_s^2 and the other σ_e^2 . The criterion Q, used to measure the efficiency of the various methods, will be the average, over all ${}^{n(n-1)/2}C_2$ possible comparisons among the $n(n-1)/2$ crosses, of this expected square error. The values of Q for methods *A*, *B* and *C* will be written Q_A , Q_B and Q_C respectively.

With all three methods of estimation, comparisons can only be made between crosses involving the inbred lines actually tested. This

is an obvious statement, but that any design not involving each inbred line at least once has $Q = \infty$ should be realized. As a result, the criterion Q cannot be used to decide on the optimal division of resources between the number of inbred lines on the one hand and the number of replicates and crosses per inbred line on the other. The number of inbred lines tested n is fixed and we need to estimate the yielding capacities of all crosses among them. A decision on the value of n would presumably involve a compromise between the accurate assessment of a relatively small number of crosses and the relatively inaccurate assessment of a large number of crosses. The more crosses tested the more intense can be the selection applied to them but the larger will be the errors involved in that selection. Problems of this kind are discussed in a series of papers by Finney [1958 a, b; 1960]. The methods he used to study mass selection in plant breeding cannot be applied directly to the present problem because of the unequal accuracies and the correlations involved in the estimation of the different crosses and because of the genetic covariance σ_g^2 between crosses involving a common line. Also, the aim may not be simply the selection of the best crosses but, for example, the selection of the best four inbred lines from which to form a four-way cross.

Two other limitations of the present approach should be noted. Firstly, we are considering only one property of the crosses, namely "yield". With more than one property, a compromise may have to be made and a design chosen that is optimal for one quantity but not for another. Secondly, we are considering all crosses among a set of inbred lines. Of more interest sometimes would be all crosses between two different sets of inbred lines that are chosen to bring together certain properties that are present in one set but not in the other.

The algebra needed to calculate the values of Q for the two methods of estimation A and B is very tedious. The comparisons among the crosses have to be divided into the three categories: sampled versus sampled, unsampled versus unsampled, and sampled versus unsampled. Using equations (3.2), (3.3) and (3.4) to sum over these categories, we find,

$$Q_A = \frac{2}{s(n-2)(n+1)} [\{2n(n-2)sa^0 + ns^2 - 2ns - n + 2\} \cdot \{\sigma_s^2 + (\sigma_e^2/r)\} + \{n^2 - 2ns - n + 2\}s\sigma_s^2]$$

and

$$Q_B = \frac{2}{s(n-2)(n+1)} [\{2n(n-2)sa^0 - n + 2\}\{\sigma_s^2 + (\sigma_e^2/r)\} + \{n^2 - 5n + 2\}s\sigma_s^2],$$

where a^0 is, as before, the diagonal element of the inverse matrix $\mathbf{A}^{-1}$. The value of Q for C, the common tester method, is

$$Q_C = \frac{4(n-1)}{(n+1)t}\left[\sigma^2_{s(t)} + (\sigma^2_e/r)\right] + 2\sigma^2_s ,$$

where t is the number of common testers, r the number of replicates, and $\sigma^2_{s(t)}$ the variance of specific combining ability among the crosses between the inbred lines and the common testers. The quantity $2\sigma^2_s$ occurs in Q_C because the estimate of each cross comparison $i \times j$ versus $l \times m$ is biased by the amount $s_{ij} - s_{lm}$. To compare A, B and C for the same total number of crosses and experimental plots, we shall assume that $t = s/2$, and that the two values of r are the same. The assumption that $\sigma^2_{s(t)} = \sigma^2_s$ has already been criticized in connection with the estimation of σ^2_g and σ^2_s . However an assumption has to be made about the ratio $\sigma^2_{s(t)}/\sigma^2_s$ and $\sigma^2_{s(t)}/\sigma^2_s = 1$ seems the most reasonable. Other values could be studied by using the general form for Q_C given above.

A comparison between the two methods A and B is immediately possible.

$$Q_A - Q_B = \frac{2n(s-2)}{(n-2)(n+1)}\left[(\sigma^2_e/r) - \sigma^2_s\right], \quad \text{for all } n \text{ and } s.$$

As is otherwise obvious, the two methods are identical when $s = 2$. If $s > 2$, A is preferred to B, i.e. $Q_A < Q_B$, if and only if $\sigma^2_e/r\sigma^2_s < 1$. If $r > 1$, a decision between A and B could be based on an estimate of $\sigma^2_e/r\sigma^2_s$ obtained from the analysis of variance of Table 1. To compare method C with methods A and B, the values of a^0 are required. The matrix $\mathbf{A}^{-1}$ can be easily determined when $s = 2$ and when $s = n - 1$ (the complete diallel cross). When $s = 2$, $\mathbf{A}^{-1}$ is generated as a circulant by the first row

$$\left(\frac{n}{4}, \frac{n}{4} - 1, \frac{n}{4} - 2, \cdots, -\frac{n}{4} + \frac{3}{2}, -\frac{n}{4} + \frac{1}{2}, -\frac{n}{4} + \frac{1}{2}, -\frac{n}{4} + \frac{3}{2}, \cdots, \frac{n}{4} - 1\right)$$

and, when $s = n - 1$, $\mathbf{A}^{-1}$ has all diagonal elements equal to $(2n - 3)/2(n - 1)(n - 2)$ and all non-diagonal elements equal to $-1/2(n - 1)(n - 2)$. Substituting $a^0 = n/4$ in Q_A and Q_B , we have that, when $s = 2$,

$$Q_A = Q_B = Q_C + \frac{(n-1)}{(n+1)(n-2)}\left[(n-2)(n-3)\{\sigma^2_s + (\sigma^2_e/r)\} - 8\sigma^2_s\right].$$

When $n = 3$, the complete diallel cross is preferred to common testers unless $\sigma_s^2 = 0$, when they are equally efficient. When $n = 5$, the partial diallel cross with $s = 2$ is preferred to using one common tester if and only if $\sigma_e^2/r\sigma_s^2 < \frac{1}{3}$. When $n \geq 7$, using one common tester is to be preferred to the partial diallel cross with $s = 2$ for all values of $\sigma_e^2/r\sigma_s^2$. Substituting $a^0 = (2n - 3)/2(n - 1)(n - 2)$ in the expression for Q_B, we have that for the complete diallel cross ($s = n - 1$)

$$Q_B + \frac{4}{(n-2)(n+1)}\{(n-3)(\sigma_e^2/r) + (3n-5)\sigma_s^2\}.$$

The complete diallel cross with estimation method B is therefore always to be preferred to using $(n - 1)/2$ common testers. The complete dialled cross with estimation method A will be even better if $\sigma_e^2/r\sigma_s^2 < 1$.

Values of a^0 for s intermediate to $s = 2$ and $s = n - 1$ are not so easy to obtain. An explicit expression for a^0 is derived in an appendix at the end of this paper. Expressions for the other elements of the inverse matrix $\mathbf{A}^{-1}$ are also derived and may be of use in obtaining variances for comparing the yielding capacities of particular crosses. The formula for a^0 is

$$a^0 = \frac{1}{n}\left\{\frac{1}{2s} + \sum_{j=1}^{n-1} \frac{\sin\frac{j\pi}{n}}{s\sin\frac{j\pi}{n} - \sin\frac{(n-s)j\pi}{n}}\right\}.$$

Values for a^0 calculated from this formula on an IBM 650 electronic computer agreed very well with known values for $s = n - 1$. The agreement for $s = 2$ and n large was not so good. The values of a^0 for $s = 2$ in Table 5 were calculated by $a^0 = n/4$. All other values were calculated electronically using the formula above. They are thought to be sufficiently accurate for our purpose but not necessarily accurate to the number of decimals shown.

Table 6 shows the values of Q_A, Q_B and Q_C for these same values of n and s. The first figure for each method is the coefficient in Q of σ_e^2/r and the second is the coefficient of σ_s^2. The bordered squares are those in which the common tester method is to be preferred to the diallel cross method. They are the squares for which s/n is small. In all other squares, the best method is A or B according as $\sigma_e^2/r\sigma_s^2$ is less than or greater than $\sigma_e^2/r\sigma_s^2 = 1$. The only exception to this is that when $s = 2$, A and B are equivalent. Also, for each value of n there are, presumably, intermediate values of s for which the decision between the diallel cross and common testers depends on the unknown value of $\sigma_e^2/r\sigma_s^2$. The examples of this in Table 6 are $n = 101$, $s = 10$; $n = 50$, $s = 7$;

TABLE 5

VALUES OF a^0 FOR VARIOUS n AND s

	s									
n	2	3	4	5	6	7	10	20	50	100
101	25.25		2.588		0.8332		0.2369	0.06515	0.02059	0.01005
100		6.031		1.365		0.5530				
75	18.75		1.969		0.6523		0.1978	0.06027	0.02034	
74		4.515		1.044		0.4387				
51	12.75		1.391		0.4836		0.1616	0.05575	0.02020	
50		3.119		0.7514		0.3326				
21	5.25		0.6739		0.2732		0.1162	0.05132		
20		1.394		0.3868		0.2007				
11	2.75		0.4248		0.2014		0.1056			
10		0.7689		0.2604		0.1589				
9	2.25		0.3721		0.1904					
8		0.6458		0.2387		0.1548				
7	1.75		0.3233		0.1833					
6		0.5139		0.2250						
5	1.25		0.2917							
4		0.4167								
3	0.75									

$n = 75, s = 74; a^0 = 0.01361; n = 9, s = 8; a^0 = 0.1339.$

$n = 10, s = 3; n = 8, s = 3$ and $n = 5, s = 2$, when the partial diallel cross is preferred to common testers if and only if $\sigma_e^2/r\sigma_s^2 < 0.25, 0.42, 0.63, 5.93$ and 0.33 respectively.

For given values of n and σ_e^2/σ_s^2, Table 6 could be used to compare the values of Q_A, Q_B and Q_C for different values of r and s. r will have to be greater than 1 if σ_e^2 is to be estimated and s greater than 2 if σ_s^2 is to be estimated.

6. *Estimation of General Combining Abilities*

The variance of the estimate of the difference between two general combining abilities using the partial diallel cross is

TABLE 6

COEFFICIENTS* OF σ_e^2/r AND σ_s^2 IN Q_A, Q_B AND Q_C FOR A RANGE OF VALUES OF n AND s

n		s = 2		3		4		5		6		7		10		20		50		100	
	A	**100.0**	**101.9**	23.9	25.8	**10.28**	**12.13**	5.46	7.26	**3.38**	**5.14**	2.29	4.01	**1.10**	**2.70**	**0.62**	**1.82**	**1.04**	**1.04**	2.00	0.00
101	B	**100.0**	**101.9**	23.9	25.8	**10.25**	**12.17**	5.40	7.32	**3.30**	**5.22**	2.19	4.11	**0.94**	**2.86**	**0.26**	**2.18**	**0.081**	**2.00**	0.040	1.96
100	C	**3.92**	**5.92**	2.61	4.61	**1.96**	**3.96**	1.57	3.57	**1.31**	**3.31**	1.12	3.12	**0.78**	**2.78**	**0.39**	**2.39**	**0.157**	**2.16**	0.078	2.08
																				s = 74	
	A	**74.0**	**75.9**	17.8	19.7	**7.82**	**9.59**	4.20	5.92	**2.69**	**4.36**	1.86	3.48	**0.99**	**2.46**	**0.72**	**1.64**	**1.38**	**0.68**	2.00	0.00
75	B	**74.0**	**75.9**	17.8	19.7	**7.77**	**9.66**	4.11	6.01	**2.57**	**4.46**	1.73	3.62	**0.78**	**2.67**	**0.24**	**2.13**	**0.080**	**1.97**	0.053	1.95
74	C	**3.89**	**5.89**	2.60	4.60	**1.95**	**3.95**	1.56	.56	**1.30**	**3.30**	1.11	3.11	**0.78**	**2.78**	**0.39**	**2.39**	**0.156**	**2.16**	0.105	2.10
	A	**50.0**	**51.8**	12.3	14.0	**5.53**	**7.21**	3.06	4.66	**2.05**	**3.57**	1.50	2.93	**0.95**	**2.15**	**0.94**	**1.34**	**2.00**	**0.00**		
51	B	**50.0**	**51.8**	12.2	14.1	**5.45**	**7.29**	2.94	4.78	**1.89**	**3.73**	1.30	3.14	**0.63**	**2.47**	**0.22**	**2.06**	**0.078**	**1.92**		
50	C	**3.85**	**5.85**	2.56	4.56	**1.92**	**3.92**	1.54	3.54	**1.28**	**3.28**	1.10	3.10	**0.77**	**2.77**	**0.38**	**2.39**	**0.15**	**2.15**		
	A	**20.0**	**21.6**	5.38	6.77	**2.75**	**3.97**	1.77	2.73	**1.43**	**2.24**	1.28	1.82	**1.24**	**1.25**	**2.00**	**0.00**				
21	B	**20.0**	**21.6**	5.28	6.88	**2.55**	**4.17**	1.45	3.05	**1.03**	**2.65**	0.75	2.35	**0.43**	**2.05**	**0.19**	**1.81**				
20	C	**3.64**	**5.64**	2.41	4.41	**1.82**	**3.84**	1.45	3.45	**1.21**	**3.21**	1.03	3.03	**0.73**	**2.75**	**0.36**	**2.36**				
	A	**10.0**	**11.3**	2.96	3.69	**1.92**	**2.37**	1.59	1.41	**1.52**	**1.16**	1.69	0.60	**2.00**	**0.00**						
11	B	**10.0**	**11.3**	2.74	3.92	**1.52**	**2.78**	0.91	2.09	**0.71**	**1.97**	0.55	1.73	**0.37**	**1.63**						
10	C	**3.33**	**5.33**	2.18	4.18	**1.67**	**3.67**	1.31	3.31	**1.11**	**3.11**	0.94	2.94	**0.67**	**2.67**						
																s = 8					
	A	**8.00**	**9.09**	2.52	2.89	**1.80**	**1.86**	1.69	0.88	**1.68**	**0.71**	2.00	0.00			**2.00**	**0.00**				
9	B	**8.00**	**9.09**	2.22	3.18	**1.29**	**2.38**	0.80	1.77	**0.65**	**1.74**	0.52	1.48			**0.46**	**1.54**				
8	C	**3.20**	**5.20**	2.07	4.07	**1.60**	**3.60**	1.24	3.24	**1.07**	**3.07**	0.89	2.89			**0.80**	**2.80**				
	A	**6.00**	**6.80**	2.10	1.81	**1.77**	**1.17**	2.00	0.00	**2.00**	**0.00**										
7	B	**6.00**	**6.80**	1.67	2.24	**1.07**	**1.87**	0.71	1.29	**0.60**	**1.40**										
6	C	**3.00**	**5.00**	1.90	3.90	**1.50**	**3.50**	1.29	3.29	**1.00**	**3.00**										
	A	**4.00**	**4.22**	2.00	0.00	**2.00**	**0.00**														
5	B	**4.00**	**4.22**	1.20	0.80	**0.89**	**1.11**														
4	C	**2.67**	**4.67**	1.60	3.60	**1.33**	**3.33**														
	A	**2.00**	**0.00**																		
3	B	**2.00**	**0.00**																		
	C	**2.00**	**4.00**																		

*Values in bold face type are for the first value of n shown for the row; remaining values are for the second value of n.

$$V(\hat{g}_i - \hat{g}_j) = 2(a^0 - a^{|i-j|})[\sigma_s^2 + (\sigma_e^2/r)].$$

Expressions for a^0 and $a^{|i-j|}$ derived in the appendix at the end of this paper could be used to evaluate this variance for each pair of values of i and j. We shall consider only the average of this variance over the $n(n-1)/2$ possible comparisons of general combining abilities. Using (3.4), this average variance is,

$$\text{Av. } V(\hat{g}_i - \hat{g}_j) = 2\left\{\frac{na^0}{n-1} - \frac{1}{2s(n-1)}\right\}[\sigma_s^2 + (\sigma_e^2/r)].$$

With $s/2$ common testers, the variance of all comparisons will be

$$V(\hat{g}_i - \hat{g}_j) = 4/s[\sigma^2_{s(t)} + (\sigma_e^2/r)].$$

Again making the big assumptions that the general combining ability of each line with the common testers is the same quantity as the general combining ability of each line with all other inbred lines and that $\sigma^2_{s(t)} = \sigma_s^2$, the average efficiency of the partial diallel cross relative to common testers in estimating general combining abilities is $E = 4(n-1)/(2nsa^0 - 1)$. We have taken the total number of crosses as well as the total number of experimental plots the same for both designs. Values of E can be calculated from Table 5 and are shown in Table 7. When $s = 2$, $E = 4/(n+1)$ and when $s = n-1$ (the complete diallel cross), $E = 2(n-2)/(n-1)$. The partial diallel cross is very inefficient compared with common testers when $s = 2$ for all values of n above 3. The complete diallel cross is always at least as efficient as common testers and tends to be twice as efficient as n becomes large. The partial diallel cross with intermediate values of s is often at least as efficient as common testers. If having slightly unequal standard errors for the different comparisons is not a serious disadvantage, the partial or complete diallel cross is a better design for estimating general combining abilities than common testers providing that a sufficient number of crosses can be made so that s/n is a reasonably large fraction. Providing $s > 2$, the standard errors for comparisons can be estimated by using the analysis of variance shown in Table 1 for the partial (and complete) diallel cross and the usual analysis of variance for common testers.

For a fixed total number of crosses, the partial diallel cross does permit more lines to be included in the experiment than does the complete diallel cross. In selecting lines for general combining ability, the resulting increase in the selection intensity will often more than compensate for the decreased accuracy with which each line is assessed.

7. *Summary*

A diallel cross among n inbred lines with no parental or reciprocal crosses involves a total of $n(n-1)/2$ crosses. Clearly, this number

TABLE 7

EFFICIENCY* OF PARTIAL DIALLEL CROSS RELATIVE TO COMMON TESTERS IN ESTIMATING GENERAL COMBINING ABILITIES

	s									
n	2	3	4	5	6	7	10	20	50	100
101, 100	**0.04**	0.11	**0.19**	0.29	**0.40**	0.51	**0.84**	**1.52**	**1.93**	**1.98**
75, 74	**0.05**	0.15	**0.25**	0.38	**0.51**	0.64	**1.00**	**1.51**	**1.95**	[s = 74, **1.97**]
51, 50	**0.08**	0.21	**0.35**	0.52	**0.68**	0.85	**1.22**	**1.77**	**1.96**	
21, 20	**0.18**	0.46	**0.71**	1.00	**1.18**	1.38	**1.68**	**1.90**		
11, 10	**0.33**	0.80	**1.10**	1.44	**1.56**	1.69	**1.80**			
9, 8	**0.40**	0.93	**1.24**	1.55	**1.64**	1.71	[s = 8, **1.75**]			
7, 6	**0.50**	1.14	**1.40**	1.60	**1.67**					
5, 4 3	**0.67** **1.00**	1.33	**1.50**							

* Values in boldface type are for the first value of n shown for the row; remaining values are for the second value of n.

increases rapidly with n. With limited facilities, this may mean that a complete diallel cross can only be made among a rather small number of inbred lines. The design presented in this paper allows a large number of inbred lines to be studied by performing only a sample of all the possible crosses among them. The efficiency of the design for the estimation of the variances of the general and specific combining abilities of the lines, for the prediction of the yielding capacities of the various crosses and for the estimation of the general combining ability of each of the lines, is discussed. The design is shown often to be more efficient than other designs that have been proposed.

8. *Acknowledgement*

We are indebted to Aaron Booker and Mary Clem for computations of a^0 values and other quantities.

APPENDIX

Inversion of Matrix A

A is a real symmetric circulant matrix and so can be inverted explicitly without much difficulty. Consider the general circulant matrix

$$\mathbf{A} = \begin{Bmatrix} a_0 & a_1 & \cdots & a_{n-1} \\ a_{n-1} & a_0 & \cdots & a_{n-2} \\ \cdots & \cdots & & \cdots \\ \cdots & \cdots & & \cdots \\ \cdots & \cdots & & \cdots \\ a_1 & a_2 & \cdots & a_0 \end{Bmatrix}.$$

Let $w_i = \exp.\ \{j(2\pi i/n)\}$, $(j = 1, 2, \cdots, n)$, be the n-th roots of unity. Then

$$\mathbf{A}\begin{bmatrix} 1 \\ w_i \\ w_i^2 \\ \cdot \\ \cdot \\ \cdot \\ w_i^{n-1} \end{bmatrix} = (a_0 + a_1 w_i + \cdots + a_{n-1} w_i^{n-1}) \begin{bmatrix} 1 \\ w_i \\ w_i^2 \\ \cdot \\ \cdot \\ \cdot \\ w_i^{n-1} \end{bmatrix}.$$

Therefore, the characteristic roots of **A** are

$$\lambda_j = a_0 + a_1 w_j + \cdots + a_{n-1} w_j^{n-1}, \qquad j = 1, 2, \cdots, n. \tag{1}$$

We can invert these equations to obtain the a's in terms of the λ's,

$$a_0 = \frac{1}{n}\left[\lambda_1 + \lambda_2 + \cdots + \lambda_n\right],$$

and

$$a_j = \frac{1}{n}\left[\lambda_1 w_{n-1}^j + \lambda_2 w_{n-2}^j + \cdots + \lambda_n\right], \qquad j = 1, 2, \cdots, n-1.$$

Now, $\mathbf{A}^{-1}$ is also a circulant matrix and has characteristic roots $1/\lambda_j (j = 1, 2, \cdots, n)$. Therefore its elements can be written

$$a^0 = \frac{1}{n}\left[\frac{1}{\lambda_1} + \frac{1}{\lambda_2} + \cdots + \frac{1}{\lambda_n}\right]$$

and

$$a^j = \frac{1}{n}\left[\frac{1}{\lambda_1} w_{n-1}^j + \frac{1}{\lambda_2} w_{n-2}^j + \cdots + \frac{1}{\lambda_n}\right], \quad j = 1, 2, \cdots, n-1. \quad (2)$$

Since the matrix **A** is real and symmetric, its characteristic roots λ_j and the elements of the inverse matrix are both real. Substituting $a_0 = s$, $a_1 = a_2 = \cdots = a_{k-1} = a_{k+s} = a_{k+s+1} = \cdots = a_{n-1} = 0$ and $a_k = a_{k+1} = \cdots = a_{k+s-1} = 1$ in (1) and taking real parts,

$$\lambda_j = s - \frac{\sin (n - s)\dfrac{j\pi}{n}}{\sin \dfrac{j\pi}{n}} \qquad j = 1, \cdots, n-1,$$

and (3)

$$\lambda_n = 2s.$$

Taking real parts on both sides of equations (2),

$$a^0 = \frac{1}{n}\left[\frac{1}{\lambda_1} + \frac{1}{\lambda_2} + \cdots + \frac{1}{\lambda_n}\right],$$

and

$$a^j = \frac{1}{n}\sum_{l=1}^{n} \frac{1}{\lambda_l} \cos \frac{j(n-l)}{n} 2\pi, \qquad j = 1, 2, \cdots, n-1. \quad (4)$$

The elements of the inverse matrix can therefore be obtained by substituting the equations (3) for the λ's into the equations (4). This method was used to compute the values of a^0 in Table 5. For small values of n, the methods available for inverting general matrices may be preferable to the special method outlined above for inverting circulant matrices.

REFERENCES

1. Cockerham, C. C. [1954]. An extension of the concept of partitioning hereditary variance for analysis of covariances among relatives when epistasis is present. *Genetics 39*, 859–82.
2. Comstock, R. E. and Robinson, H. F. [1948]. The components of genetic variance

in populations of biparental progenies and their use in estimating the average degree of dominance. *Biometrics 4*, 254–66.

3. Comstock, R. E. and Robinson, H. F. [1952]. Estimation of average dominance of genes. Chapter 30 of *Heterosis*, edited by J. W. Gowen. Iowa State College Press, Ames.
4. Finney, D. J. [1958a]. Statistical problems of plant selection. *Bull. de l'Inst. Internat. de Stat. 36* (3), 242–68.
5. Finney, D. J. [1958b]. Plant selection for yield improvement. *Euphytica 7*, 83–106.
6. Finney, D. J. [1960]. A simple example of the external economy of varietal selection. *Bull. de l'Inst. Internat. de Stat., 37* (3), 91–106.
7. Fisher, R. A. [1918]. The correlation between relatives on the supposition of Mendelian inheritance. *Trans. Roy. Soc. Edinb. 52*, 399–433.
8. Griffing, B. [1956a]. A generalized treatment of the use of diallel crosses in quantitative inheritance. *Heredity 10*, 31–50.
9. Griffing, B. [1956b]. Concept of general and specific combining ability in relation to diallel crossing systems. *Aust. J. Biol. Sci. 9*, 463–93.
10. Hayman, B. I. [1954a]. The analysis of variance of diallel tables. *Biometrics 10*, 235–44.
11. Hayman, B. I. [1954b]. The theory and analysis of diallel crosses, *Genetics 39*, 789–809.
12. Jensen, S. [1959]. *Combining ability of unselected inbred lines of corn from incomplete diallel and top-cross tests.* Unpublished Ph.D. Thesis, Iowa State University.
13. Jinks, J. L. and Hayman, B. I. [1953]. The analysis of diallel crosses *Maize Genetics Coop. News Letter 27*, 48–54.
14. Kempthorne, O. [1953]. A class of experimental designs using blocks of two plots. *Ann. Math. Stat. 24*, 76–84.
15. Kempthorne, O. [1956]. The theory of the diallel cross. *Genetics 41*, 451–59.
16. Kempthorne, O. [1957]. *An introduction to genetic statistics.* John Wiley and Sons, New York.
17. Sprague, G. F. and Tatum, L. A. [1942]. General vs. specific combining ability in single crosses of corn. *J. Amer. Soc. Agron. 34*, 923–32.
18. Zoellner, J. A. and Kempthorne, O. [1954]. Incomplete block designs with blocks of two plots. *Agricultural Experiment Station Research Bulletin 418*, Iowa State College, Ames, Iowa.

SENSORY TESTING BY TRIPLE COMPARISONS

G. T. Park
T. Hedley and Co. Ltd., Gosforth, Newcastle upon Tyne, England.

1. INTRODUCTION

In sensory test work, the number of samples a judge can assess at any one time is often limited by sensory fatigue and confusion of stimuli. This necessitates the use of incomplete block designs. Further, the difficulty of obtaining satisfactory quantitative measures of treatment effects usually entails assessment by ranking. Thus there is need for a rank analysis of incomplete block designs of a versatility equivalent to that available for variables satisfying the conditions of analysis of variance. In addition, any such analysis should provide the experimenter with a check as to whether he can regard the treatments concerned as lying on a linear scale of acceptability. For, with sensory assessment, cases can and do arise where the relations between calculated average scale positions disagree with individual direct comparisons.

These needs have been met in the case of blocks of size two, by the various paired comparison procedures developed in recent years [2, 4–6; 13, 14, 18, 17]. Although this is undoubtedly the most important case, instances arise where blocks of size three or more are a practical proposition and hold out the possibility of increased efficiency.

Unfortunately, little work seems to have been done in the general case of blocks of size k. The result published by Durbin [7] allows an overall significance test of treatment differences on balanced incomplete block ranking experiments, whilst the rankit method of Fisher [10] has been used by some workers, but no generally accepted method (of adequate versatility) has been available for blocks of size other than two. The method recently published by Pendergrass and Bradley [3, 15] which offers a comprehensive treatment of the case $k = 3$ is therefore of immediate interest.

This paper is to assist evaluation of the Bradley-Pendergrass method by contrasting results from triple comparison testing with results obtained using paired comparisons. Two sets of data are discussed. One set is from margarine cocktail stick taste tests, where sensory fatigue is known to be marked. The other set is from toilet soap sniff tests, where sensory fatigue is much less of a problem.

2. BRADLEY-PENDERGRASS TRIPLE COMPARISON ANALYSIS

The members of each of the $\binom{t}{3}$ triplets formed by the set of t treatments are ranked in order of acceptability by n judges.[1] Analysis proceeds as follows.

The Model

The model used is an extension of that successfully developed by Bradley and Terry for paired comparisons, [2, 4–6, 8, 11]. It is supposed that the treatments $T_1, \cdots, T_t$ are associated with parameters $\pi_1, \cdots, \pi_t$ which satisfy the following conditions:

$$\pi_i \geq 0, \qquad \sum_{i=1}^{t} \pi_i = 1,$$

$$P(T_i > T_j > T_k) = \pi_i^2 \pi_j / \Delta_{ijk},$$

$$\Delta_{ijk} = \pi_i^2(\pi_j + \pi_k) + \pi_j^2(\pi_i + \pi_k) + \pi_k^2(\pi_i + \pi_j),$$

$$i \neq j \neq k; \qquad i, j, k = 1, \cdots, t,$$

where $P(T_i > T_j > T_k)$ represents the probability that in the block containing treatments T_i, T_j, T_k, treatment i is rated top, treatment j central and treatment k bottom, in acceptability.

Maximum Likelihood Estimators of Parameters π_i

The maximum likelihood estimators $p_1, \cdots, p_t$ of the parameters $\pi_1, \cdots, \pi_t$ are given by the equations below:

$$a_i/p_i = n \sum_{\substack{j<k \\ i,k \neq i}} \{[2p_i(p_j + p_k) + p_j^2 + p_k^2]/D_{ijk}\},$$

$$\sum_i p_i = 1, \qquad i, j, k = 1, \cdots, t$$

where a_i = twice the total number of first places + the total number of second places given to T_i and

$$D_{ijk} = p_i^2(p_j + p_k) + p_j^2(p_i + p_k) + p_k^2(p_i + p_j).$$

In most cases these can be solved by one or two iterations using as initial estimates the values of p_i found from the quadratic equations:

$$p_i^2[a_i(2t^2 - 6t + 6) - n(2t^4 - 11t^3 + 22t^2 - 19t + 6)]$$
$$+ p_i[a_i(2t - 6) - n(t^3 - 4t^2 + 5t - 2)] + 2a_i = 0, \qquad i = 1, \cdots, t.$$

[1] Extension of the analysis to the general case of n_{ijk} rankings on the triplet T_i, T_j, T_k, $n_{ijk} \geq 0$, $i < j < k$, $i, j, k = 1 \cdots t$, is readily made and results quoted by the authors in [15].

Variance—Covariance Matrix of Estimators p_i

Expressions for the variances and covariances are presented by Bradley and Pendergrass for large n. These are particularly awkward to compute, however, and the authors provide an approximation which should suffice for practical purposes. If $\pi_1 \cdots \pi_t$ are near equality, then

$$\text{Var}\,(\sqrt{n}p_i) = 3(t-1)/t^4(t-2), \qquad i = 1, \cdots, t,$$

$$\text{Cov}\,(\sqrt{n}p_i\,,\ \sqrt{n}p_j) = -3/t^4(t-2), \quad i \neq j, \quad i, j = 1, \cdots, t.$$

Goodness of Fit

The model can be tested by comparing observed frequencies f_{ijk} with expected frequencies f'_{ijk}. If f_{ijk} = number of times treatments T_i, T_j, T_k were ranked in that order of acceptability, and

$$f'_{ijk} = np_i^2p_j/D_{ijk}\,, \qquad i \neq j \neq k, \qquad i, j, k = 1, \cdots, t,$$

then

$$\sum_{\substack{i \neq j \neq k \\ i,j,k}} \frac{(f_{ijk} - f'_{ijk})^2}{f'_{ijk}}$$

is distributed as chi-square with $\left[5\binom{t}{3} - \left(t-1\right)\right]$ degrees of freedom for large n.

Scaling of Treatments

If the data fit the model, then the authors suggest that the treatments be regarded as lying at points $\ln p_i$ on a linear scale of acceptability. This is by analogy with the paired comparison model of Bradley and Terry. In that case, the parameters π_i can be shown to be related to a metric on which the mean response for treatment T_i is $\ln \pi_i$.

Overall Tests of Main Effects and Interaction

An overall test for main effects is available if required. For large n,

$$Z = 2n\binom{t}{3} \ln 6 + 2 \sum_i a_i \ln p_i - 2n \sum_{i<j<k} \ln D_{ijk}$$

is distributed as chi-square with $(t-1)$ degrees of freedom. Further, if the experiment has been repeated over a number of different groups of judges or in different circumstances, interaction can be tested using the result that: $Z^c - Z$ is distributed as $\chi^2(g-1)(t-1)$ where $Z^c = \sum_{u=1}^{g} Z_u$, g = no. of groups, $Z_u = Z$ above, computed for uth group, $u = 1 \cdots g$.

Relative Efficiency of Paired Comparisons and Triple Comparisons

Under the assumption that differences observed by judges between the three component pairs of each triplet are the same as those which would be observed in the equivalent paired comparisons, then the efficiency of triple comparisons to paired comparisons is 9/4 in terms of incomplete block rankings (or $\frac{3}{2}$ in terms of treatment replications, as three assessments are made per triplet ranking against two per pair ranking).

3. PAIRED COMPARISON ANALYSES

As the various paired comparison procedures give closely similar results [1, 12], any one of them could be used as the basis for assessment of the triple comparison technique. However, that of Bradley and Terry is an obvious choice because of its relationship to the triple comparison method. In addition, the results from Scheffé's method [17] are given, as results from this method had to be obtained for routine reporting of the paired comparison legs. It must be expected, of course, that the results of these two analyses will agree more closely with each other than with the triple comparison results, as they are derived from the same basic data, whereas the triple comparison data is distinct and independent.

TABLE I

RESULTS OF TEST I

Statistic	Paired Comparisons		Triple Comparisons
	Scheffé	B-Terry	B-Pendergrass
Response Ratings[2]	α_i	$\ln p_i$	$\ln p_i$
Product A	−0.55	−0.42	−0.06
Product B	0.22	0.28	−0.34
Product C	0.36	0.43	0.40
Product D	0.45	0.28	0.50
Product E	−0.48	−0.57	−0.50
S.E. of Diff.	±0.48	±0.52	±0.33
Goodness of Fit	$\chi^2(6) = 2.3$	$\chi^2(6) = 5.7$	$\chi^2(46) = 47.5$
Treatment Diffs.	$\chi^2(4) = 8.2$	$\chi^2(4) = 5.7$	$\chi^2(4) = 13.8$

4. MARGARINE COCKTAIL STICK TASTE TESTS

It was found that guidance on the likely results of full-scale consumer research studies could be obtained by taste testing small cylinders of margarine from cocktail sticks. Two such tests were done for routine purposes by paired comparisons and for experimental purposes by triple comparisons.

Test I. In Test I, judges examined five margarines A–E. Each of 60 judges made one paired comparison to give 3 rankings per ordered pair. Each of a further 120 judges made one triple comparison to give 2 rankings per ordered triplet.

Test II. In Test II, judges examined three margarines L–N. Each of 138 judges made one paired comparison to give 23 rankings per ordered pair. Each of a further 126 judges made one triple comparison to give 21 rankings per ordered triplet.

Paired comparison and triple comparison ratings are seen to be in close agreement on Test II. On Test I agreement is not so good, there being an indication that the methods give different measures of the relative acceptability of Products *A* and *B*. The goodness of fit chi-squares show that both of the paired comparison models and the triple comparison model gave acceptable fits to their respective data.

On Test II, the relative sensitivity of the three procedures on a per ranking basis is indicated by the treatment chi-squares as they stand, for the number of pair rankings approximated the number of

TABLE 2

RESULTS OF TEST II

Statistic	Paired Comparisons		Triple Comparisons
	Scheffé	B-Terry	B-Pendergrass
Response Ratings[2]	α_i	$\ln p_i$	$\ln p_i$
Product L	−0.45	−0.38	−0.44
Product M	0.55	0.62	0.56
Product N	−0.10	−0.21	−0.12
S.E. Of Diff.	±0.46	±0.58	±0.41
Goodness of Fit	$\chi^2(1) = 0.0$	$\chi^2(1) = 0.2$	$\chi^2(3) = 1.3$
Treatment Diffs.	$\chi^2(2) = 4.8$	$\chi^2(2) = 3.2$	$\chi^2(2) = 6.0$

[2]All sets of response ratings have been adjusted to total zero and range one.

triplet rankings. In the case of Test I, it is necessary to double the paired comparison chi-squares before comparing them with the triple comparison chi-square, as the number of pair rankings was only one half the number of triplet rankings. In each case, there is no indication that triple comparisons were more efficient than paired comparisons. It is of interest that application of Durbin's chi-square test of treatment differences to the triple comparison data from Tests I and II gave values close to those found under the Bradley-Pendergrass analysis ($\chi^2(4) = 13.4$, $\chi^2(2) = 6.0$).

5. TOILET SOAP SNIFF TESTS

Two toilet soap sniff tests were done for routine purposes by paired comparisons and for experimental purposes by triple comparisons. *Test III* examined four toilet soap perfumes A–D. Each of 132 judges made one paired comparison to give 11 rankings per ordered pair.

Each of a further 131 judges made one triple comparison. Although the design arranged for rankings to be divided evenly over the 24 ordered triplets, an error in arranging record sheets resulted in the number of rankings varying as follows: (Order of presentation within

	Triplet			
	ABC	ABD	ACD	BCD
No. of rankings	51	17	11	52

triplet remained in approximate balance). Analysis of this test was therefore carried out using the extension of the method to the general case when the number of rankings differs between triplets.

Test IV examined six toilet soap perfumes L–Q. Each of 180 judges made two successive paired comparisons (involving four different products). As there was no evidence that results from pairs sniffed first differed from those for pairs sniffed second, the results for all pairs were combined to give a total of 360 rankings, i.e. 12 rankings per ordered pair. Each of a further 120 judges made one triple comparison to give one ranking per ordered triplet.

On both tests, paired comparison and triple comparison ratings are in agreement within the limits of error.

On Test III, the paired comparison models fitted the paired comparison data, but the triple comparison model was not a satisfactory fit to the triple comparison data (goodness of fit chi-square significant

TABLE 3

RESULTS OF TEST III

Statistic	Paired Comparisons		Triple Comparisons
	Scheffé	B-Terry	B-Pendergrass
Response Ratings	α_i	$\ln p_i$	$\ln p_i$
Product A	0.30	0.23	0.30
Product B	0.44	0.46	0.55
Product C	−0.56	−0.54	−0.40
Product D	−0.18	−0.15	−0.45
S.E. of Diff.[3]	±0.42	±0.50	±0.31
Goodness of Fit	$\chi^2(3) = 0.3$	$\chi^2(3) = 0.2$	$\chi^2(17) = 40.1$
Treatment Diffs.	$\chi^2(3) = 7.2$	$\chi^2(3) = 4.5$	$\chi^2(3) = 16.5$

[3]An average value is given for the S.E. of difference of triple comparison ratings as this figure varies between pairs of products due to differing numbers of rankings per triplet.

TABLE 4

RESULTS OF TEST IV

Statistic	Paired Comparisons		Triple Comparisons
	Scheffé	B-Terry	B-Pendergrass
Response Ratings	α_i	$\ln p_i$	$\ln p_i$
Product L	0.30	0.37	−0.04
Product M	−0.55	−0.45	−0.36
Product N	0.45	0.30	0.16
Product O	−0.55	−0.63	−0.55
Product P	0.13	0.19	0.34
Product Q	0.22	0.22	0.45
S.E. of Diff.	±0.38	±0.29	±0.21
Goodness of Fit	$\chi^2(10) = 5.7$	$\chi^2(10) = 6.5$	$\chi^2(95) = 76.7$
Treatment Diffs.	$\chi^2(5) = 13.5$	$\chi^2(5) = 18.5$	$\chi^2(5) = 32.8$

at 1 percent level). On Test IV, all three models fitted their respective data.

Using the treatment difference chi-squares as before to compare sensitivity on a per ranking basis, it appears that, on both tests, considerably greater sensitivity was achieved with triple comparisons than with paired comparisons. (Note that on Test IV the number of triplet rankings was only one-third the number of pair rankings so that the triple comparison chi-square requires multiplication by three before contrasting with the paired comparison chi-squares). Application of Durbin's test of treatment differences to the triple comparison data from the balanced test, Test IV, gave $\chi^2(5) = 31.5$ in good agreement with that found under the Bradley-Pendergrass analysis.

6. ANALYSIS OF VARIANCE-TEST III

As the extension of the triple comparison analysis to the case of varying numbers per triplet was not available at the time, the triple comparison results from Test III were analysed by least squares [using dummy variables [16]]. As shown below, the results agreed very closely with those subsequently found by the triple comparison method.

	Products				S.E. of Diff.	Treatment $\chi^2(3)$
	A	B	C	D		
Triple Comparison	0.30	0.55	−0.40	−0.45	±0.31	16.5
Least Squares	0.32	0.54	−0.40	−0.46	±0.31	17.1

7. DISCUSSION AND SUMMARY

These results suggest that the model on which the triple comparison analysis is based will prove appropriate to at least a proportion of sensory tests. Further, that in those cases where sensory fatigue is not marked, it may afford a valuable gain in efficiency over paired comparisons. It is pointed out, however, that each of the experiments described above compared products which were basically similar and which differed essentially in only one dimension.

In the author's experience, paired comparison tests involving treatment differences of this nature rarely give results inconsistent with the various paired comparison models. Lack of fit tends to occur when treatments differ importantly in more than one dimension or when marked judge × treatment interactions are likely to be present. As the triple comparison model is likely to be more vulnerable than the

paired comparison models to the complications arising on these latter types of test, it may prove of limited applicability where treatments differ in kind rather than in degree.

ACKNOWLEDGMENT

The author is indebted to R. A. Bradley for commenting on the first draft of this paper and to Miss Anne Thirlwall for assistance with the computation.

REFERENCES

[1] Bliss, C. I., Greenwood, Mary L., and White, Edna S. [1956]. A rankit analysis of paired comparisons for measuring the effect of sprays on flavour. *Biometrics 12*, 381–403.

[2] Bradley, R. A. and Terry, M. E. [1952]. Rank analysis of incomplete block designs I. The method of paired comparisons. *Biometrika 39*, 324–45.

[3] Bradley, R. A. and Terry, M. E. [1952]. The rank analysis of incomplete block designs II. The method for blocks of size three. *Appendix A, Bi-annual Report No. 4, Va. Agr. Exp. Sta.*

[4] Bradley, R. A. [1954]. Rank analysis of incomplete block designs II. Additional tables for the method of paired comparisons. *Biometrika 41*, 502–37.

[5] Bradley, R. A. [1954]. Incomplete block rank analysis. On the appropriateness of the model for a method of paired comparisons. *Biometrics 10*, 375–90.

[6] Bradley, R. A. [1955]. Rank analysis of incomplete block designs III. Some large-sample results on estimation and power for a method of paired comparisons. *Biometrika 42*, 450–70.

[7] Durbin, J. [1951]. Incomplete blocks in ranking experiments. *Brit. Jour. Psych. 4*, 85–90.

[8] Dykstra, O. [1956]. A note on the rank analysis of incomplete block designs—Applications beyond the scope of existing tables. *Biometrics 12*, 301–06.

[9] Dykstra, O. [1958]. Factorial experimentation in Scheffé's analysis of variance for paired comparisons. *J.A. S.A.* 53, 529–42.

[10] Fisher, R. A. and Yates, F. [1953]. *Statistical Tables for Biological, Agricultural and Medical Research.* Fourth Edition, Oliver & Boyd, Edinburgh, 25, 76.

[11] Hopkins, J. W. [1954]. Incomplete block rank analysis: some taste test results. *Biometrics 10*, 391–99.

[12] Jackson, J. E. and Fleckenstein, Mary [1957]. An evaluation of some statistical techniques used in the analysis of paired comparison data. *Biometrics 13*, 51–64.

[13] Mosteller, F. [1951 a]. Remarks on the method of paired comparisons I. The least squares solution assuming equal standard deviations and equal correlations. *Psychometrika 16*, 3–9.

[14] Mosteller, F. [1951 b]. Remarks on the method of paired comparisons II. The effect of an aberrant standard deviation when equal standard deviations and equal correlations are assumed III. A test of significance for paired comparisons when equal standard deviations and equal correlations are assumed. *Psychometrika 16*, 203–18.

[15] Pendergrass, R. N. and Bradley, R. A. [1960]. Ranking in triple comparisons. *Contributions to Probability and Statistics*, edited by I Olkin et. al., Stanford University Press, Stanford, California.
[16] Quenouille, M. H. [1950]. *Introductory Statistics*, 148–54.
[17] Scheffé, H. [1952]. An analysis of variance for paired comparisons. *J.A.S.A. 47*, 381–400.
[18] Thurstone, L. L. [1927]. Psychophysical analysis. *Amer. J. Psychol 38*, 368.

RAPID CHI-SQUARE TEST OF SIGNIFICANCE FOR THREE-PART RATIOS[1]

CHARLES E. GATES AND BENJAMIN H. BEARD[2]
University of Minnesota, St. Paul, Minnesota, U.S.A.

INTRODUCTION

In the study of qualitative characters, the number of individuals distributed among mutually exclusive categories is commonly compared with a distribution predicated on genetic theory. If the observed numbers in the various classes deviate from the theoretical, it is desirable to be able to reject the hypothetical distribution at a determinable probability level. A statistic normally used in ascertaining this goodness-of-fit is chi-square. Frequently, the number of such comparisons may be large and the computations necessary for each segregating family become laborious. In these instances an accurate, fast technique would prove useful. This paper presents a rapid, graphical procedure for testing goodness-of-fit to ratios with three classes. These graphical chi-square tests of significance are particularly appropriate for testing lack of fit to three-part ratios for those organisms where it is feasible to take random samples of some constant number of individuals from segregating progenies. However, by constructing a series of graphs, it is practical to use this method with variable sample size if the number of uses is sufficiently large. Graphs to be used with the 1 : 2 : 1 ratio have been drawn for a variety of values of n and have been used successfully in the field for determining lack of fit.

Although the justification for, as well as the illustrated examples, have been drawn from genetics, uses in other fields can be visualized. One particular application outside the field of genetics could be in evaluating the results of a sample survey where a response can be categorized into three mutually exclusive classes and one is interested in ascertaining the goodness-of-fit to a 1 : 1 : 1 ratio.

ILLUSTRATION

Suppose the numbers of individuals falling in three mutually exclusive classes are x_1 , x_2 and x_3 , where

$$x_1 + x_2 + x_3 = n. \tag{1}$$

[1]Paper No. 4346, Scientific Journal Series, Minnesota Agricultural Experiment Station.
[2]Crops Research Division, Agricultural Research Service, U. S. Department of Agriculture.

The critical region for a particular level of significance can be defined by an ellipse drawn on rectangular coordinate paper with x_2 and x_1 (or x_3) the ordinate and the abscissa, respectively. Figure 1 was prepared for the 1 : 2 : 1 ratio, the three ellipses indicating demarcation between the .10, .05 and .01 levels of significance.

To illustrate the use of this graph, suppose the segregates from individual F_1 plants from a cross of two barley varieties differing in number of rows of kernels on the spike are classified and counted. In a cross of this kind (two-row VV × six-row vv) the heterozygote (Vv) is phenotypically distinguishable from either homozygote and, if conditioned by one gene, the F_2 should segregate in a 1 : 2 : 1 ratio. Consider the following examples wherein x_1 , x_2 and x_3 refer respectively to the numbers in the VV, Vv and vv classes:

Example	x_1	x_2	x_3	n
1	20	16	14	50
2	16	22	12	50
3	7	33	10	50
4	15	33	2	50

Locating the points of intersection of x_1 and x_2 (or x_3 and x_2), as measured along the abscissa and ordinate, respectively, the probabilities from the χ^2 statistic can be immediately ascertained as $.05 > P > .01$, $P > .10$, $.10 > P > .05$ and $P < .01$ for examples 1, 2, 3 and 4, respectively.

THEORETICAL DEVELOPMENT FOR AN ARBITRARY THREE-PART RATIO

Genetic theory or other considerations predicate that three classes corresponding to those indicated by (1) should have the proportions π_1 , π_2 and π_3 , respectively. It can readily be seen that

$$E(x_i) = n\pi_i \ , \quad \text{Var}\,(x_i) = n\pi_i(1 - \pi_i), \quad \text{Cov}\,(x_i \ , x_j) = -n\pi_i\pi_j \ . \tag{2}$$

It is well known (see, e.g., Cochran [1952]) that

$$[x_1 - n\pi_1 \ x_2 - n\pi_2]\begin{bmatrix} n\pi_1(1 - \pi_1) & -n\pi_1\pi_2 \\ -n\pi_1\pi_2 & n\pi_2(1 - \pi_2) \end{bmatrix}^{-1}\begin{bmatrix} x_1 - n\pi_1 \\ x_2 - n\pi_2 \end{bmatrix} \tag{3}$$

will be distributed asymptotically as chi-square with 2 degrees of freedom. Expression (3) may be re-written as

$$[(x_1 - n\pi_1)^2\pi_2(1 - \pi_2) + 2\pi_1\pi_2(x_1 - n\pi_1)(x_2 - n\pi_2) + (x_2 - n\pi_2)^2\pi_1(1 - \pi_1)]/(n\pi_1\pi_2\pi_3). \tag{4}$$

Now (4) is algebraically identical to the chi-square goodness-of-fit statistic

$$\chi^2_{2,\alpha} = \sum_{i=1}^{3} [(x_i - n\pi_i)^2]/n\pi_i = \left[\sum_{i=1}^{3} x_i^2/n\pi_i\right] - n \tag{5}$$

indicating, incidentally, why the latter has its well known distribution.

Equation (5) can therefore be seen to represent an ellipse, not in standard form, in the x_1 , x_2 plane, with the center at

$$(n\pi_1 , n\pi_2). \tag{6}$$

The acceptance region of the usual goodness-of-fit test is the region inside this ellipse.

In order to plot ellipse (5), it is necessary to determine the slopes of the minor and major axis as well as their lengths. If the x_1 , x_2 axes are rotated so that ellipse (5) will be in standard form, the slope of the rotated x_1 axis is

$$m = R + \sqrt{R^2 + 1} \tag{7}$$

where

$$R = (\pi_1 - \pi_2)\pi_3/2\pi_1\pi_2 .$$

Rotating the ellipse, (5) can be written

$$[A(x_1 - n\pi_1)^2/\chi^2_{2,\alpha}] + [C(x_2 - n\pi_2)^2/\chi^2_{2,\alpha}] = 1 \tag{8}$$

where

$$A = [\pi_2(1 - \pi_2) + 2m\pi_1\pi_2 + \pi_1(1 - \pi_1)m^2]/n\pi_1\pi_2\pi_3(1 + m^2), \tag{9}$$

$$C = [\pi_2(1 - \pi_2)m^2 - 2m\pi_1\pi_2 + \pi_1(1 - \pi_1)]/n\pi_1\pi_2\pi_3(1 + m^2). \tag{10}$$

The semi-diameters of the minor and major axes are therefore

$$\chi_{2,\alpha}/\sqrt{A}, \qquad \chi_{2,\alpha}/\sqrt{C}, \tag{11}$$

respectively, where $A, C > 0$.

The slope of the rotated axis, as indicated by (7), is independent of sample size, n, as well as significance level. On the other hand, the location of the center of an ellipse, as shown by (6), is a function of n, but is independent of significance level. The lengths of the major and minor axes, given by (11), are a function of both sample size and significance level. Hence a family of ellipses with a common center and slope can be constructed for given n by varying significance level. In using these ellipses, it is recommended that the expected number in any class be the customary minimum of 5 or 10 because the distribution of (3) is only asymptotically distributed as χ^2.

CONSTRUCTION OF ELLIPSES FOR THE 1:2:1 RATIO

Substitution of 1/4, 1/2 and 1/4 for π_1 , π_2 and π_3 , respectively, yields the following basic quantities:

$$R = -1/4, \quad m = 0.780776, \quad 1 + m^2 = 1.609612,$$

$$A = 11.123103/n, \qquad C = 2.876897/n.$$

Making use of (6), the center of the 1 : 2 : 1 ellipse, is located at

$$(n/4, n/2). \tag{12}$$

Substituting the basic quantities in (11), one-half the lengths of the minor and major axes are found to be

$$.2998\sqrt{n}\chi_{2,\alpha} \quad \text{and} \quad .5896\sqrt{n}\chi_{2,\alpha}\,, \tag{13}$$

respectively. An ellipse can be readily constructed for each n in the range of interest using a Dietzgen ellipsograph or various simple, but accurate construction techniques such as the concentric-circle method for major and minor diameters. French and Vierck [1958], a recent reference, demonstrate various methods of ellipse construction.

To illustrate the procedure more specifically, consider ellipses for $n = 50$, where the critical regions are defined by the 0.10, 0.05 and 0.01 levels of significance. Making use of (12) and (13), one obtains These ellipses were plotted in Figure 1.

Significance level	Slope of rotated x_1 axis	Location of center		Semi-diameter length	
		x_1	x_2	x_1	x_2
.10	.781	12.50	25.00	4.55	8.94
.05	.781	12.50	25.00	5.20	10.22
.01	.781	12.50	25.00	6.43	12.65

CONSTRUCTION OF ELLIPSES FOR OTHER THREE-PART RATIOS

Calculation of ellipses for defining critical regions for other three-part ratios can be carried out in analogous fashion to the ellipses for the 1 : 2 : 1 ratio. To eliminate the necessity of deriving intermediate calculations, the slopes, center, and lengths of major and minor diameters for some of the more common three-part genetic ratios and the 1 : 1 : 1 ratio are given in Table 1. These quantities are expressed in terms of arbitrary n and significance level. To obtain figures for graph-

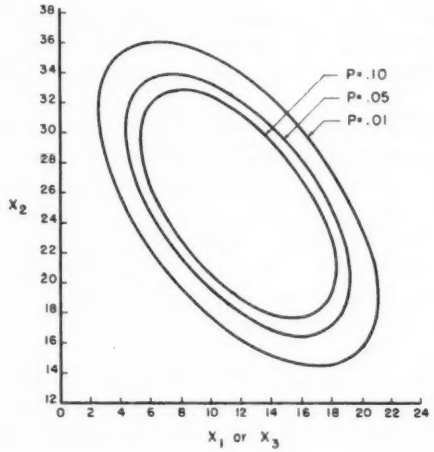

FIGURE 1

χ^2 TEST FOR A 1:2:1 RATIO

x_1, x_2 and x_3 are the observed numbers of progenies in the VV, Vv and vv classes, respectively. Intersection of x_1 and x_3 (or x_3 and x_2) outside an ellipse indicates significance at the $P = .01$, .05 and .10 levels of significance for sample size $n = 50$.

TABLE 1

SLOPES, CENTERS AND SEMI-DIAMETERS OF ELLIPSES DEFINING CRITICAL REGIONS FOR VARIOUS THREE-PART RATIOS.

Genetic Ratio*	Slope of minor axis of ellipse relative to x_1 axis	Center		Length of semi-diameters	
		x_1	x_2	x_1	x_2
3:9:4	0.650	$3n/16$	$9n/16$	0.289θ‡	0.561θ
6:9:1	0.973	$3n/8$	$9n/16$	0.171θ	0.672θ
3:9:3	0.721	$n/5$	$3n/5$	0.271θ	0.571θ
3:9:1	0.895	$3n/13$	$9n/13$	0.186θ	0.597θ
3:12:1	0.883	$3n/16$	$3n/4$	0.168θ	0.558θ
1:14:1	0.638	$n/16$	$7n/8$	0.154θ	0.380θ
1:2:1	0.781	$n/4$	$n/2$	0.300θ	0.590θ
1:1:1†	1.000	$n/3$	$n/3$	0.408θ	0.408θ

*x_2 was written as the largest component of each ratio. This is somewhat at variance from usual genetic terminology.

†non-genetic

‡$\theta = \sqrt{n}\,\chi_{2,\alpha}$

ing, it is necessary only to substitute n and $\chi^2_{2,\alpha}$ and to plot the ellipses as previously described. For three-part ratios not shown, it is necessary to substitute appropriate values for π_1 , π_2 , and π_3 in (6), (7), (9) and (10).

It would be possible to rewrite (4) in terms of the proportion $p_i = x_i/n$, resulting in corresponding changes in the above results. For a given genetic ratio and significance level, a family of ellipses could then be constructed with common center, the larger samples leading to the smaller ellipses. Such graphs would prove in some respects more convenient to use than the procedure outlined above. However, this convenience would in some instances, be offset by the requirement of the calculation of two of the p_i in order to determine the level of significance. The decision as to the better procedure depends on whether it is more desirable to have a minimum of separate graphs, but require supplementary computations, or to have more graphs with no computation necessary. Separate ellipses in either instance need not be constructed for each n as ellipses for n of similar size will be similar.

We are indebted to a referee for pointing out a more elegant approach to the derivation of the characteristics of the chi-square ellipses and to Donald Richter for helpful comments.

REFERENCES

Cochran, W. G. [1952]. The χ^2 test of goodness-of-fit. *Ann. Math. Stat. 23*, 315–45.
French, T. E. and Vierck, C. J. [1958]. *Graphic Science*. McGraw Hill, New York.

GENERATING UNBIASED RATIO AND REGRESSION ESTIMATORS

W. H. WILLIAMS

Bell Telephone Laboratories, Incorporated
Murray Hill, New Jersey, U. S. A.

1. INTRODUCTION

Information collected on a concomitant variate is often used in finite sampling theory to create more precise estimators of population characteristics. This supplementary information is obtained in addition to the characteristic under study and some aspects of it may be derived from sources other than the sample itself. It may be either qualitative or quantitative. For example, suppose that the variate under consideration in a sample survey is the number of dairy cattle per farm y and that at the time of the survey the number of grazing acres per farm x is also obtained. It may then be known from census data that the total number of grazing acres in the entire area is $N\mu_X$ and the mean per farm is μ_X. Analytically, we have a random sample of n pairs (y_i, x_i), $i = 1, \cdots, n$, from a population of size N and the population x-mean is known exactly. The problem is to estimate the population mean μ_Y.

A general class of estimators designed to utilize this supplementary information includes ratio and regression estimators. These estimators are described in textbooks on the subject, see for example Cochran [1953]. Additional developments have been presented by Hartley and Ross [1954], Nicto [1958] and Robson [1957].

The two classical ratio estimators are the ratio of means estimator $\tilde{y} = (\bar{y}/\bar{x})\mu_X$ and the mean of ratios estimator $\hat{y} = \mu_X \sum_{i=1}^{n} r_i/n$ where $\bar{y}$ and $\bar{x}$ are sample means and $r_i = y_i/x_i$. It is well known that these estimators are biased. The usual regression estimator is obtained by evaluating the least squares line of best fit $y = \bar{y} + b(x - \bar{x})$ at the point μ_X giving $\hat{y}_b = \bar{y} + b(\mu_X - \bar{x})$ as a regression estimator of μ_Y. This estimator is biased if the assumption of a linear model is not valid.

The generation of some exactly unbiased ratio and regression estimators is discussed in this paper. Specifically, we classify an estimator as of the regression type if it is invariant under location and scale changes in x and if it undergoes the same location and scale changes

as the y variate. A ratio estimator has analogous properties but for scale changes only.

2. DERIVATION OF UNBIASED ESTIMATORS FOR SIMPLE RANDOM SAMPLING

To generate unbiased estimators, consider the following sampling procedure. At step one, select with equal probability one of all possible splits of the population into s mutually exclusive groups of size[1] n/k, i.e., $N = sn/k$. At the second stage, select randomly without replacement k of the groups from the total number of groups s of that particular split of the population. This gives a sample of size n.

Now consider the conditional distribution for a particular set of s groups. Attached to each of these groups there are characteristics[2] $\bar{y}^{(i)}$, $\bar{x}^{(i)}$, $b^{(i)}$, $i = 1, \cdots, s$, where $\bar{y}^{(i)}$ and $\bar{x}^{(i)}$ are means of the n/k units in the group and $b^{(i)}$ is as yet an unspecified function of the y and x of that group. For a given split and a random selection of groups, the expectations of $\bar{y}^i$ and $\bar{x}^i$, $i = 1, \cdots, k$, are μ_Y and μ_X respectively; that is, they are conditionally unbiased. Furthermore,

$$\left(1 - \frac{k}{s}\right) \frac{1}{k(k-1)} \sum_{i=1}^{k} (b^i - \bar{b})(\bar{x}^i - \bar{x}) \tag{1}$$

is an unbiased estimator of Cov $(\bar{b}, \bar{x})$ where $\bar{b} = \sum_{i=1}^{k} b^i/k$.

Hence if $g = \bar{y} + \bar{b}(\mu_X - \bar{x})$ then $E(g) = \mu_Y - \text{Cov}\ (\bar{b}, \bar{x})$ and

$$T_k = \bar{y} + \bar{b}(\mu_X - \bar{x}) + \left(1 - \frac{n}{N}\right) \frac{1}{k(k-1)} \sum_{i=1}^{k} (b^i - \bar{b})(\bar{x}^i - \bar{x}) \tag{2}$$

is a conditionally unbiased estimator of μ_Y. It is then unbiased unconditionally.

This approach is valid for any defined form of the coefficient $b^{(i)}$; T_k will remain unbiased. If $b^{(i)}$ has a form which is invariant under linear x and y transformation (say least squares form) then T_k is classified as a regression estimator. If $b^{(i)} = \bar{y}^{(i)}/\bar{x}^{(i)}$ (say), then T_k falls into the class of a ratio estimator.

This procedure is used to generate the unbiased estimators; in practice a simple random sample would be drawn and to compute T_k it would be split randomly into groups, see Section 7 for an example. The latter operation is equivalent to the generating procedure which allows a particular split-sample to arise in $\binom{s}{k}(N - n)!/[(n/k)!]^{s-k}$

[1]It is assumed that this relationship is true in terms of integers.

[2]Superscripts will be used to specify the groups. They will be used with parentheses when the reference is to the entire population of s groups and without parentheses when referring to the sample of k groups.

ways while splitting the simple random sample allows a particular split-sample to arise in only one way. The unbiasedness is preserved by either procedure.

The argument is easily generalized to p auxiliary variates.

3. SPECIFIC ILLUSTRATIONS IN SIMPLE RANDOM SAMPLING

A form of interest is

$$b^{(i)} = \frac{\sum_{j=1}^{n/k} (y_j - \bar{y}^{(i)})(x_j - \bar{x}^{(i)})}{\sum_{j=1}^{n/k} (x_j - \bar{x}^{(i)})^2}, \qquad i = 1, \cdots, s, \tag{3}$$

the least squares slope form.

In this case T_k bears much resemblance to $\hat{y}_b$ and might be thought of as possessing an additional component which is required to compensate for possible bias in $\hat{y}_b$. This is not exactly true of course, because the first two terms of T_k are not exactly the two terms of $\hat{y}_b$. However, in this case (3), it seems natural to make some remarks on the efficiency of T_k.

The variance of T_k depends very much on the form of the $b^{(i)}$ coefficients. In fact, until the form of $b^{(i)}$ is specified little can be said about the variance of T_k. One can imagine choices which would lead to poor efficiency indeed. However, T_k in this case has coefficients in the least squares slope form and it is natural to ask how it compares with $\hat{y}_b$ when a linear model is assumed, for then $\hat{y}_b$ has optimum variance properties. But with this assumption, $\hat{y}_b$ also possesses unbiasedness and the advantage of T_k is unbiasedness in situations in which $\hat{y}_b$ is not unbiased. However, one would like the efficiency of T_k to compare favorably even in this linear model case. So by assuming a linear model and a normal x-distribution, it is easily found that $V(\hat{y}_b)/V(T_k) = (n-2)(n-6)/(n-3)(n-4)$, $k = 2$ and $n > 6$. This expression is less than one but approaches one as n gets larger and, for example, when $n = 15, 25$ is equal to 0.89 and 0.95. Thus we see that one does not lose all the efficiency brought about by the use of an auxiliary variate and that $[V(T_k) - V(\hat{y}_b)]/V(\hat{y}_b)$ is $O(n^{-1})$.

Furthermore, the role of k also depends upon the choice of the $b^{(i)}$. For example, in the special case of the previous paragraph, if the number of groups is regarded as variable, $V(\hat{y}_b)/V(T_k)$ will be found to have a maximum at $k = \sqrt{n/3}$. Thus for this form of the $b^{(i)}$, the optimum number of groups is $\sqrt{n/3}$. Other forms of the $b^{(i)}$ would yield other results.

Another possible choice is

$$b^{(i)} = \sum_{i=1}^{n/k} y_i x_i \Big/ \sum_{i=1}^{n/k} x_i^2 .$$

In this form T_k is a ratio estimator and it is unbiased even if the linear relationship of y and x does not pass through the origin. But characteristically the variance will be inflated by such a relationship.

Next, if $b^{(i)} = \bar{y}^{(i)}/\bar{x}^{(i)} = r^{(i)}$, T_k will reduce to the form

$$T_k = \bar{r}\mu_X + \frac{Nk - n}{N(k-1)}(\bar{y} - \bar{r}\bar{x}) \tag{4}$$

where $\bar{b}$ is denoted $\bar{r}$. It will be noted that when $k = n$, $T_k = y'$, the unbiased ratio estimator presented by Hartley and Ross [1954]. The efficiency of this form of T_k has been examined in detail by Goodman and Hartley [1958] and Robson [1957]. Robson presents an exact variance formula for finite populations.

Finally, consider $b^{(i)} = r^{(i)} = (k/n)\sum_{i=1}^{n/k} r_i$, $r_i = y_i/x_i$, then $\bar{b} = \bar{r} = \sum_{i=1}^{n} r_i/n$ which does not depend upon the particular split of the population. Now if, after substitution of this form into T_k , the estimator is averaged over all possible splits of the sample into groups of size n/k it will be found that the result is again the Hartley-Ross unbiased ratio estimator. This averaging process is indicated by a star, i.e., T_k^* .

Other forms could, of course, be considered.

4. STRATIFIED SAMPLING

Since a bias may be magnified relative to the standard deviation, stratified sampling may perhaps be regarded as the most important application of unbiased estimators. Their separate use within strata requires exact knowledge of the population strata means but is straightforward. We now develop a combined stratified estimator.

Consider L strata of size N_t , $t = 1, \cdots, L$ with $\sum_{t=1}^{L} N_t = N$, and again consider the sampling in two stages. At the first stage select with equal probability one of the possible splits of each stratum into s groups of size n_t/k, $t = 1, \cdots, L$. Then $N_t = sn_t/k$. At the second stage select k groups with equal probability and without replacement from each of the strata, giving a sample of size n_t in the t-th stratum, $\sum_{t=1}^{L} n_t = n$.

For a given split and a random selection of groups

$$\bar{y}_{st}^{i} = \sum_{t=1}^{L} (N_t/N)\bar{y}_t^i \quad \text{and} \quad \bar{x}_{st}^{i} = \sum_{t=1}^{L} (N_t/N)\bar{x}_t^i$$

are unbiased estimators of μ_Y and μ_X respectively, where $\bar{y}_t^i$ and $\bar{x}_t^i$ denote means of the i-th group in the t-th stratum. Also we can consider

a coefficient $b_{st}^{(i)}$ which is as yet unspecified in form but utilizes the set of elements in the i-th group of all strata. For example,

$$b_{st}^{(i)} = \frac{\sum_{t=1}^{L}\sum_{j=1}^{n_t/k}(y_{tj} - \bar{y}_t^{(i)})(x_{tj} - \bar{x}_t^{(i)})}{\sum_{t=1}^{L}\sum_{j=1}^{n_t/k}(x_{tj} - \bar{x}_t^{(i)})^2}, \qquad i = 1, \cdots, s \tag{5}$$

is an over-all slope estimator.

Next we note that

$$\bar{y}_{st} = \sum_{t=1}^{L}(N_t/N)\bar{y}_t = \sum_{i=1}^{k}\bar{y}_{st}^{i}/k \tag{6}$$

where $\bar{y}_t$ is the mean of the n_t observations in the t-th stratum (similarly for $\bar{x}_{st}$) and finally that a conditionally unbiased estimator of Cov $(\bar{b}_{st}, \bar{x}_{st})$ is given by

$$(1 - n/N)\frac{1}{k(k-1)}\sum_{i=1}^{k}(\bar{x}_{st}^{i} - \bar{x}_{st})(b_{st}^{i} - \bar{b}_{st}).$$

Consequently, if $g = \bar{y}_{st} + \bar{b}_{st}(\mu_X - \bar{x}_{st})$ then $E(g) = \mu_Y -$ Cov $(\bar{b}_{st}, \bar{x}_{st})$ and therefore

$$T_{k(st)} = \bar{y}_{st} + \bar{b}_{st}(\mu_X - \bar{x}_{st}) + \left(1 - \frac{n}{N}\right)\frac{1}{k(k-1)}\sum_{i=1}^{k}(\bar{x}_{st}^{i} - \bar{x}_{st})(b_{st}^{i} - \bar{b}_{st}) \tag{7}$$

is a combined stratified unbiased estimator of μ_Y. Note that since $N_t = sn_t/k$, $k/s = n/N$. Nieto [1958] discussed the efficiency of the estimator (7) (for sampling with replacement) in detail.

Again the generalization to p auxiliary variates is straightforward.

As a specific illustration consider the case in which

$$b_{st}^{(i)} = \bar{y}_{st}^{(i)}/\bar{x}_{st}^{(i)} = r_{st}^{(i)}.$$

Then $T_{k(st)}$ reduces to

$$T_{k(st)} = \bar{r}_{st}\mu_X + \frac{Nk - n}{N(k-1)}(\bar{y}_{st} - \bar{r}_{st}\bar{x}_{st}). \tag{8}$$

In the special case that $N_t = \bar{N}$, $n_t = \bar{n}$ for all t and $k = \bar{n}$, $s = \bar{N}$ then

$$T_{k(st)} = \bar{r}_{st}\mu_X + \frac{(\bar{N} - 1)\bar{n}}{(\bar{n} - 1)\bar{N}}(\bar{y}_{st} - \bar{r}_{st}\bar{x}_{st}), \tag{9}$$

which is a generalized Hartley-Ross estimator.

Finally, we again consider an averaging of T_k over all possible splits of the sample into groups of size n_t/k, $t = 1, \cdots, L$. For this, the

coefficient is taken in the form $b_{st}^{(i)} = r_{st}^{(i)} = \sum_{t=1}^{L} (N_t/N) r_t^{(i)}$ where $r_t^{(i)} = (k/n_t) \sum_{i=1}^{n_t/k} (y_i/x_i)$. Therefore,

$$\bar{r}_{st} = \sum_{i=1}^{k} r_{st}^{i} \Big/ k = \sum_{t=1}^{L} (N_t/N) \sum_{i=1}^{n_t} (y_i/x_i) \Big/ n_t = \sum_{t=1}^{L} (N_t/N) \bar{r}_t$$

and some algebraic reduction will show that $T_{k(st)}$ averaged over all possible splits is equal to

$$T_{k(st)}^{*} = \bar{r}_{st}\mu_X + (\bar{y}_{st} - \bar{r}_{st}\bar{x}_{st}) + \left(1 - \frac{n}{N}\right) \sum_{t=1}^{L} \frac{N_t^2}{N^2} \frac{(\bar{y}_t - \bar{r}_t\bar{x}_t)}{(n_t - 1)}, \qquad (10)$$

which does not quite reduce to a form similar in appearance to Equation (8) and the Hartley-Ross estimator.

As before other selections of coefficients will yield other unbiased estimators.

5. MULTISTAGE SAMPLING

We consider a population with N primaries of equal size $\bar{M}$ and the following sampling scheme. First select n primaries from the N available with equal probability with or without replacement. Then select with equal probability one of the splits of each of the primaries into s groups of size $\bar{m}/k$. Then with equal probability and without replacement draw k of the groups so that the sample size is $\bar{m}$ in each selected primary.

Consider now the conditional distribution for a fixed set of primaries and a fixed split of the primaries into s groups each. Then by Section 4, Equation (11) is an unbiased estimator of $\bar{Y}_n$, the population mean of the n selected primaries.

$$T_{k(M)} = \bar{y} + b(\mu_X - \bar{x}) + \left(1 - \frac{\bar{m}}{\bar{M}}\right) \frac{1}{k(k-1)} \sum_{i=1}^{k} (b^i - b)(\bar{x}^i - \bar{x}) \qquad (11)$$

where

$$\bar{y}^i = (k/n\bar{m}) \sum_{t=1}^{n} \sum_{j=1}^{\bar{m}/k} y_{tj}^{i}, \qquad \bar{y} = (1/k) \sum_{i=1}^{k} \bar{y}^i = (1/n\bar{m}) \sum_{t=1}^{n} \sum_{j=1}^{\bar{m}} y_{tj}$$

and similarly for x. The coefficient $b^{(i)}$ is again arbitrary in form.

Finally, the expectation of $T_{k(M)}$ over all possible primary selections is the average of $\bar{Y}_n$ over all possible primary selections; this is μ_Y and $T_{k(M)}$ is unbiased in multi-stage sampling.

Again the selection of the coefficients yields estimators of different types. For example, an unbiased ratio estimator of the Hartley-Ross type generalized to multistage sampling can be obtained.

6. VARIANCE ESTIMATION

It is interesting to notice that the same two-stage sampling scheme can be used to form an estimate of the variance of T_k. First assume a negligible n/N (or k/s) and a fixed set of uncorrelated groups. T_k can now be written

$$T_k = \bar{y} - \frac{1}{k(k-1)} \sum_{i \neq j}^{k} b^i(\bar{x}^i - \mu_X) \tag{12}$$

and its conditional variance can be expressed in terms of the variances and covariances of the components in (12). Since T_k is conditionally unbiased this variance has expectation equal to the over-all variance. Substitution of unbiased estimators for each of the terms of the variance (plus some terms of zero expectation) yields (13) as an unbiased estimator of the variance of T_k.

$$v(T_k) = T_k^2 - \frac{1}{k(k-1)} \sum_{i \neq j}^{k} \bar{y}^i \bar{y}^j. \tag{13}$$

Although this procedure is unbiased it can be subject to high sampling error, particularly for small k.

TABLE 1

A SIMPLE EXAMPLE OF THE ESTIMATORS

Sample Number	Pairs in Sample	$\bar{y}$	T_2			T_2^*
			Split 1	Split 2	Split 3	
1	$P_1P_2P_3P_4$	3.500	7.167	6.667	6.500	6.778
2	$P_1P_2P_3P_5$	5.250	8.917	8.250	7.917	8.361
3	$P_1P_2P_3P_6$	7.500	11.000	10.167	9.667	10.278
4	$P_1P_2P_4P_6$	8.750	11.917	10.250	9.917	10.694
5	$P_1P_2P_4P_5$	6.500	10.000	8.667	8.500	9.056
6	$P_2P_3P_4P_5$	7.500	8.167	7.667	7.500	7.778
7	$P_2P_3P_5P_6$	11.500	10.000	8.667	8.500	9.056
8	$P_2P_3P_4P_6$	9.750	9.417	8.750	8.417	8.861
9	$P_1P_3P_4P_5$	7.250	9.417	8.750	8.417	8.861
10	$P_1P_3P_4P_6$	9.500	11.000	10.167	9.500	10.222
11	$P_1P_3P_5P_6$	11.250	11.917	10.250	9.917	10.694
12	$P_3P_4P_5P_6$	13.500	7.167	6.667	6.500	6.778
13	$P_1P_4P_5P_6$	12.500	11.000	10.167	9.667	10.278
14	$P_2P_4P_5P_6$	12.750	8.917	8.250	7.917	8.361
15	$P_1P_2P_5P_6$	10.500	13.167	10.667	10.500	11.444

7. NUMERICAL ILLUSTRATION

To illustrate T_k a small population consisting of the six pairs $P_i = (y_i , x_i)$, $i = 1, 2, \cdots , 6$, with $y = x^2$ and $x_i = 0, 1, 2, \cdots , 5$ was completely examined. Table 1 presents the values of $\bar{y}$, T_2 and T_2^* [using Equation (3)] for all possible samples of size four. For T_2 , each of the possible samples was split into two groups of two in all possible ways, and the value of T_2 was computed for each. The three distinct values of T_2 for each sample are presented in the table. The numbering of the splits within a sample is of course arbitrary. It is readily verified that the average value of each of $\bar{y}$, T_2 and T_2^* is the population mean $\mu_Y = 9.167$. Furthermore, the exact population variances of $\bar{y}$, T_2 and T_2^* are 7.914, 2.281 and 1.886 respectively.

As a second example, the six pairs (y_i , x_i) were taken as follows: (0,2), (1, 3), (2, 5), (4, 9), (8, 14), (9, 15). All possible samples of size four were drawn and for each sample $\bar{y}$, y', $\hat{y}_b$, T_k (for all possible splits) and T_k^* were computed. A summary of the computations is presented in Table 2.

TABLE 2

ILLUSTRATION OF RELATIVE EFFICIENCIES

	Estimator				
	$\bar{y}$	y'	$\hat{y}_b$	T_k	T_k^*
Expectation	3.937	4.000	3.961	4.000	4.000
Bias	−0.063	0.000	−0.039	0.000	0.000
Variance	0.120	0.233	0.027	0.033	0.022
Mean Square Error	0.124	0.233	0.029	0.033	0.022

REFERENCES

Cochran, W. G., [1953]. *Sampling Techniques.* New York, John Wiley and Sons.

Goodman, L. A., and Hartley, H. O., [1958]. The precision of unbiased ratio-type estimators. *J. Amer. Stat. Assoc.* 53. 491–508.

Hartley, H. O., and Ross, A., [1954]. Unbiased ratio estimators. *Nature 174*, 270–271.

Nieto, J., [1958]. *Unbiased ratio estimators in stratified sampling.* Unpublished M.S. Thesis, Iowa State University, Ames, Iowa.

Robson, D. S., [1957]. Application of multivariate polykays to the theory of unbiased ratio-type estimation. *J. Amer. Stat. Assoc.* 52. 511–522.

A BIOMETRIC THEORY OF MIDDLE AND LONG DISTANCE TRACK RECORDS[1]

MALCOLM E. TURNER AND ELEANOR D. CAMPBELL
Medical College of Virginia, Richmond 19, Virginia, U. S. A.

1. *Introductory Remarks.*

It has been often observed that a plot of average velocity against distance or time for world track records suggests a continuous monotonic underlying relationship. Several attempts to discern the nature of this relationship empirically by seeking linearizing transformations have been made with varying degrees of success. The purposes of this paper are to describe an attempt to explain the observed curve by a biometric theory and to consider the problem of estimating unknown parameters occurring in the theory. The theory is applied to world record data of 1959 and records which are below par are pointed out. The development is based on the work of A. V. Hill [1927].

2. *Biometric Theory.*

Consider the situation in which a runner runs a given distance x in a certain time t at a constant speed $\dot{x}$, this constant speed not necessarily being the runner's maximum speed. We further define:

y, the total energy (measured as volume of oxygen) excess over resting required to run the given distance with the given constant speed,

$\dot{x}_0$, the greatest speed for which the energy (oxygen) required is not greater than the energy available for an indefinitely long period of time. Thus, at speed $\dot{x}_0$ the energy required is equal to the maximum amount of oxygen which can be brought into the body from the atmosphere,

r, the maximum excess (over resting) rate at which oxygen can be brought into the body from the atmosphere,

D, the maximum tolerable oxygen debt beyond which complete exhaustion is manifest.

The total excess energy y is a function of the constant speed $\dot{x}$ and the time of the race t. We assume that the rate of consumption of energy (oxygen) $\partial y/\partial t$ is a function of $\dot{x}$ but not of t. Thus,

[1]Read before the spring meeting of ENAR, April 1960, and based in part on a Ph.D. thesis supervised by Robert J. Monroe and H. L. Lucas, Jr., North Carolina State College, 1959.

$$\partial y/\partial t = f(\dot{x}) \text{ and } y = f(\dot{x})t.$$

The constant of integration is taken to be zero since no excess energy is required to run a race of zero duration.

The function $f(\dot{x})$ is of unknown form; however, we expect it to have at least the following characteristics:

(1) $f(\dot{x})$ is a non-decreasing function of $\dot{x}$.
(2) $f(\dot{x})$ has lower-order derivatives which vanish at the equilibrium velocity $\dot{x}_0$,
(3) $f(\dot{x}_0) = r$.

The first characteristic has been experimentally demonstrated by Hill [1927]. The second characteristic, although not experimentally demonstrated, seems plausible by analogy with other physiological equilibrium phenomena. It often happens that in the neighborhood of a point of equilibrium (here there is an equilibrium between the energy being consumed and the oxygen being brought into the cells) that compensatory physiological mechanisms come into play. It may be that circulatory changes or viscosity changes (due to temperature changes) provide such compensation in the present case. Whether or not these mechanisms are plausible, it seems appropriate to require characteristic (2) in the absence of experimental fact since one is hard pressed to find any physiological example not possessing this characteristic. Characteristic (3) is implied by the definition of $\dot{x}_0$ and r.

Perhaps the simplest function possessing these characteristics is the following:

$$f(\dot{x}) = r + b(\dot{x} - \dot{x}_0)^k, \qquad \dot{x} \geq \dot{x}_0 . \tag{1}$$

Then we have

$$y = rt + b(\dot{x} - \dot{x}_0)^k t. \tag{2}$$

Equation (2) can be used to describe the experimental data shown in Figure 11 of Hill [1927]. Hill measured directly the excess oxygen required by an athlete to run for given periods of time at constant speeds. This experiment, of course, does not correspond to actual contest conditions, but does serve to illustrate the relation (2), which was not given by Hill. The meaning of the relation is clear. If a runner runs at his own asymptotic speed $\dot{x}_0$, then the excess energy required is obviously proportional to the time run where the proportionality constant is r; but, if he runs faster than $\dot{x}_0$, then the energy required will be increased monotonically with the divergence of his speed from the asymptotic speed.

Now, the speed cannot be increased indefinitely since there is an upper bound on energy available. This upper bound is given by

$$y_{\max} = rt + D. \tag{3}$$

For all but the shorter races the time lost in attaining optimum speed is negligible and it can be seen that the best time for a race can be obtained by maintaining a nearly constant speed throughout the race. It may be supposed that records are broken when a runner with a great amount of available energy is able to select that constant speed which equates the maximum available energy to the required energy. In this case

$$rt + D = rt + b(\dot{x} - \dot{x}_0)^k t\,, \tag{4}$$

which upon cancelling rt from both sides and rearranging gives an equation relating speed to time run. Thus,

$$\dot{x} = \dot{x}_0 + (bt/D)^{-1/k}. \tag{5}$$

The constants $\dot{x}_0$, D, b, and k are constants pertaining to the individual runner. D is the maximum oxygen debt the runner can tolerate and b and k relate to the mechanical efficiency of the runner's body, i.e., to the amount of work required to overcome the viscosity of the muscles. $\dot{x}_0$ is a function of the rate at which oxygen can be gotten to the tissues. Relation (5) could be used to study the progress of an individual runner in his period of training; however, our interest at present is in characterizing the world record curve. Although $\dot{x}_0$, D, b, and k are personal parameters, it may be supposed that, for world record holders, $\dot{x}_0$ and D approach the human maximum for all and that b and k are similarly close to the human minimum for all and that therefore it is appropriate to treat $\dot{x}_0$, D, b, and k as constant parameters for all record holders as a first approximation. Equation (5) is a special case of the single process law described by Turner [1959].

3. *Estimation Theory.*

For convenience we change the notation. Let:

$$Y = \dot{x}, \qquad \delta = -1/k,$$

$$\alpha = \dot{x}_0\,, \qquad X = t,$$

$$\beta = (b/D)^{-1/k} = (D/b)^{1/k}.$$

Then (5) may be rewritten:

$$Y = \alpha + \beta X^{\delta}. \tag{6}$$

Now, (6) cannot be expected to hold exactly for world records but it may be considered to be the limiting curve which is approached by the successive records in each event. Thus, for the actual observed records, we may write

$$Y = \alpha + \beta X^{\delta}\epsilon', \tag{7}$$

where $\epsilon' \leq 1$. Perhaps, the limiting curve (6) itself is not absolutely stable since the human species is possibly changing in its athletic capacities due to nutritional and genetic effects. However, these changes may be regarded to take place slowly as compared with the rate at which records are broken.

At any point in time we will assume as a first approximation that the distribution of world records below its asymptotic value can be described by the *truncated hyperbolic distribution*:[2]

$$f(\epsilon') = 1/\rho\epsilon', \qquad e^{-\rho} \leq \epsilon' \leq 1. \tag{8}$$

If we substract α from both sides of (7) and take logarithms we obtain the linear form:

$$\log (Y - \alpha) = \log \beta + \delta \log X + \epsilon, \tag{9}$$

where $\epsilon = \log \epsilon'$. Now it will be seen that the additive error ϵ follows the *rectangular distribution*:

$$f(\epsilon) = 1/\rho, \qquad -\rho \leq \epsilon \leq 0. \tag{10}$$

Thus, the problem of estimation is reduced to a well known one: to fit a straight line when the errors have a rectangular distribution. Maximum likelihood estimates are readily found by minimizing the maximum error (see Turner, [1960]). The procedure is as follows:

(a) Choose a trial estimate of α and then plot $\log (Y - \alpha)$ against $\log X$.

(b) The line with minimum maximum error will either (i) pass through one point such that all other points fall below and such that two points will have the same maximum error, or (ii) pass through two points such that all other points fall below and such that one point will have the maximum error.

(c) In case (i) the line will be parallel (and hence have the same slope) as a line passing through the two points with maximum error. The intercept is then found by passing a line with the same slope through the upper point. In case (ii) both slope and

[2]The truncated hyperbolic distribution was chosen as being, perhaps, the simplest distribution producing high contact on the right and having a finite upper limit. Since the upper limit is the object of estimation, the lower portions of the distribution are relatively irrelevant.

TABLE I

WORLD TRACK RECORDS AND ESTIMATED LIMITING VALUES

Distance		Time			Speed in m./sec.	
x Metric	English	t Record*	$\hat{t}$ Estimated	$\hat{t} - t$ Difference	$\dot{x}$ Record	$\hat{\dot{x}}$ Estimated
400.0 m.		45.2″	45.2″	0.0″	8.850	8.850
402.3 m.	440 yd.	45.7″	45.6″	0.1″	8.804	8.822
800.0 m.		1′ 45.7″	1′ 44.3″	1.4″	7.569	7.670
804.7 m.	880 yd.	1′ 46.8″	1′ 45.1″	1.7″	7.534	7.657
1,000.0 m.		2′ 18.1″	2′ 15.5″	2.6″	7.241	7.380
1,500.0 m.		3′ 38.1″	3′ 35.4″	2.7″	6.878	6.964
1,609.4 m.	1 mi.	3′ 54.4″	3′ 53.1″	1.3″	6.866	6.904
2,000.0 m.		5′ 2.2″	4′ 58.0″	4.2″	6.618	6.711
3,000.0 m.		7′ 52.8″	7′ 47.2″	5.6″	6.345	6.421
3,218.7 m.	2 mi.	8′ 32.0″	8′ 24.9″	7.1″	6.287	6.375
4,828.0 m.	3 mi.	13′ 10.8″	13′ 4.9″	5.9″	6.105	6.151
5,000.0 m.		13′ 35.0″	13′ 34.7″	0.3″	6.135	6.137
9,656.0 m.	6 mi.	27′ 43.8″	27′ 29.0″	14.8″	5.804	5.856
10,000.0 m.		28′ 30.4″	28′ 30.4″	0.0″	5.847	5.847
Asymptotic Speed						5.103

*Source: *Information Please Almanac 1960* [1959].

intercept are found for the line passing through the two upper points. For either case the maximum error is readily ascertained. The choice of case (i) or (ii) is quickly made graphically and other determinations are made analytically.

(d) The above three steps are repeated for other choices of α and then, from the relation between minimum maximum error and α, one interpolates for the maximum likelihood estimate of α. Repeating the above three steps one final time with the optimum value of α leads to maximum likelihood estimates for β, δ, and ρ.

The above procedure illustrates a general theorem concerning minimum maximum error approximation by polynomial equations due to Vallée-Poussin [1911].

4. *Fit of Theory to World Track Records.*

The 1959 World records for all events between 400 and 10,000 meters (including English distances of 440 yards to six miles) were used to fit the theory. These data are given in Table I. Shorter distances and longer distances bring into play mechanisms unaccounted for in the theory, such as delay due to starting, muscle stiffness, blisters, nutritional requirements, etc.

Performing steps (a), (b), and (c) we obtain the following results:

$\hat{\alpha}$	$\hat{\rho}$ Minimized Maximum Error
5.08	0.03249
5.09	0.03190
5.10	0.03130
5.11	0.03143
5.12	0.03183

Parabolic interpolation between the middle three values in the above table yields the value $\hat{\alpha} = 5.103$. Refitting the straight line gives the estimates $\hat{\delta} = -0.4451$, $\hat{\beta} = 20.44$, and $\hat{\rho} = 0.03115$. We then have the estimated limiting curve.

$$\hat{\hat{x}} = 5.103 + 20.44t^{-0.4451}, \tag{11}$$

where the estimated asymptotic value of the speed is $\hat{\hat{x}}_0 = 5.103$. The estimated limiting speeds have been calculated according to equation

(11) and are shown in the last column of Table I beside the actual record speeds. Both record and estimated speeds are also shown in Figure 1. Although no very long races have been included in the analysis it is interesting to compare the speed for 30,000 $m.$, $\dot{x} = 5.26$, with the estimated asymptotic speed, $\hat{\dot{x}}_0 = 5.10$.

Estimated times $\hat{t}$ may be readily found by dividing the distance x by the estimated speeds $\hat{\dot{x}}$. These values are shown in column (4) of Table I and the differences between these estimated limiting times and the actual record times are given in column (5) of the same table. The

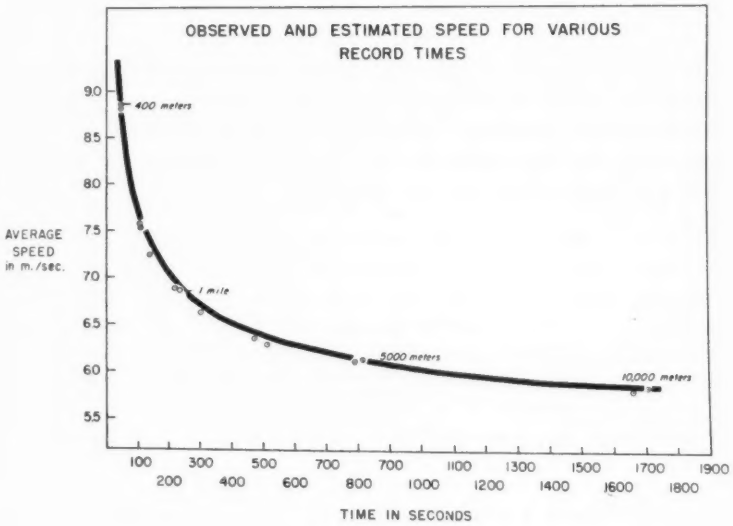

FIGURE 1

OBSERVED AND ESTIMATED SPEEDS FOR VARIOUS RECORD TIMES

extent of the agreement is almost startling. The largest discrepancy is seen to be just 14.8 seconds for the six-mile run. The time for the mile run appears to be relatively good, along with the 400, 5,000 and 10,000 meter races. With the exception of the mile run the "English" distances tend to lag behind the corresponding "metric" distances. Perhaps, this fact reflects nothing more than the popularity of the various distances, but one is tempted to speculate that large improvements in those races which lag behind may be forthcoming.

REFERENCES

Hill, A. V. [1927]. *Muscular Movement in Man.* McGraw-Hill Book Co., Inc., New York.

Information Please Almanac 1960. [1959]. McGraw-Hill Book Co., Inc., New York.

Turner, Malcolm E. [1959]. *The Single Process Law: A Study in Nonlinear Regression,* Dissertation No. 59-6571 (North Carolina State College). University Microfilms, Inc., Ann Arbor, Mich.

Turner, Malcolm E. [1960]. Note: On heuristic estimation methods. *Biometrics 16,* 299–300.

Vallée-Poussin, C. J. de la. [1911]. Sur la méthode de l'approximation minimum. *Annales de la Soc. Sc. de Bruxelles 35,* B: 1.

ADDENDUM

Since this analysis was completed Sweden's Dan Waern lowered the 1000 m. record to 2′ 17.8″ and Australia's Herb Elliott lowered the 1500 m. record to 3′ 36.0″. These are both above the estimated optimum times for these distances and hence the analysis is not affected. However, the fit is substantially improved.

SMALL SAMPLE BEHAVIOR OF SLOPE ESTIMATORS IN A LINEAR FUNCTIONAL RELATION[1]

MARTIN DORFF
University of Maine, Orono, Maine, U. S. A.
AND
JOHN GURLAND[2]
Mathematics Research Center, U. S. Army, University of Wisconsin Madison, Wisconsin, U. S. A.

1. INTRODUCTION

This paper is concerned with estimating the slope parameter in a linear functional relation between two variables which are not observable since both are subject to error. Such problems occur in many contexts. As a matter of fact, in virtually all applications of fitting a functional relation errors of observation occur in all variables. Very frequently we can neglect errors in the independent variables in comparison with the errors in the dependent variable, but clearly we cannot always do so. In particular, it is often desirable to use a relatively inexpensive technique in place of a very costly one, where the results given by the two, apart from errors, are linearly related. For example, the evaluation of rocket grains by static test techniques is relatively inexpensive compared to the procedure of utilizing dynamic or flight testing.

The following linear relation is assumed between the variables X, Y.

$$Y = \alpha + \beta X.$$

Neither X nor Y can be observed, but only x and y, where

$$x = X + e$$
$$y = Y + f;$$

e and f are independent random variables with expectation zero representing errors of observations; X and Y are the expected values of x and y respectively. When for each $X_i(Y_i)$ we have a single observation $x_i(y_i)$, we shall say that we have unreplicated observations on that variable; otherwise we have replicated observations.

[1]Preparation of this paper for publication was supported in part by contract No. DA-11-022-ORD-2059; research supported in part by Air Force Office of Scientific Research, Contract No. AF 49(639)43.

[2]On leave from Iowa State University of Science and Technology, Ames, Iowa, U. S. A.

In a previous paper [3] we considered the large sample properties of various estimators of the parameters α, β. In the present paper we consider the small sample behavior of these slope estimators in order to compare them. First however, we take up the small sample behavior of slope estimators which do not require replicated observations; such estimators were discussed only briefly in our earlier paper.

We have chosen to examine the bias and the mean square error of the various estimators, partly because of the innate interest which these properties hold and partly because they seemed most amenable to attack. In all the cases considered the square of the bias is small compared with the mean square error; consequently our findings concerning the latter are essentially valid for the variance as well.

2. ESTIMATORS WHICH DO NOT REQUIRE REPLICATION

The problem of estimating the parameters of a linear functional relation when both variables are in error and there is no replication has long been regarded as an intractable one. In 1940 Wald [10] suggested that, when the number of observations n is even, one can divide the observations into two equal groups and take as an estimator of β

$$b_W = \Big(\sum_{\frac{1}{2}n+1}^{n} y_i - \sum_{1}^{\frac{1}{2}n} y_i\Big) \Big/ \Big(\sum_{\frac{1}{2}n+1}^{n} x_i - \sum_{1}^{\frac{1}{2}n} x_i\Big).$$

He showed that b_W is a consistent estimator of β provided that the partition of observations into two groups can be carried out independently of the errors and provided that limit inferior of

$$| (X_1 + \cdots + X_{\frac{1}{2}n}) - (X_{\frac{1}{2}n+1} + \cdots + X_n) | /n$$

is positive. If the errors in the x_i are small enough so that partitioning the observations according to the magnitudes of the x_i yields the same two groups as obtained by partitioning the observations according to the magnitudes of the X_i, the first condition is satisfied. The second condition guarantees that the expectation of the denominator does not vanish.

Nair and Shrivastava [8], Bartlett [1], and others [4], [9], have shown that, for certain spacings of the X_i, more efficient estimators of β can be achieved by dividing the observations into three groups and omitting the middle group. In particular, Bartlett suggested that, when the number of observations is divisible by three, one should take as an estimator of β

$$b_B = \Big(\sum_{\frac{2}{3}n+1}^{n} y_i - \sum_{1}^{\frac{1}{3}n} y_i\Big) \Big/ \Big(\sum_{\frac{2}{3}n+1}^{n} x_i - \sum_{1}^{\frac{1}{3}n} x_i\Big).$$

This is a consistent estimator of β under the obvious modification of

Wald's conditions. Bartlett showed that b_B is more efficient than b_W when the X_i are equally spaced.

Housner and Brennan [6] suggest that the estimator

$$b_H = \sum (i - \bar{i})y_i / \sum (i - \bar{i})x_i ,$$

where i now denotes the order number when the x's are ordered according to increasing magnitude, will often be more efficient than any of the foregoing. In fact they give empirical results to substantiate the belief that b_H is more efficient than b_W when the X_i are equally-spaced. It should be realized, however, that consistency of this estimator rests upon more stringent assumptions than those made by Wald and the other writers mentioned above. One must now assume the entire ordering according to the magnitudes of the x_i to be identical with the ordering according to the X_i .

It is clear that all of the estimators mentioned above are members of the class $b_L = \sum w_i y_i / \sum w_i x_i$, where i denotes the order number when the x's are ordered according to increasing magnitude and $\sum w_i = 0$. One requires this restriction on the w_i in order to ensure that b_L be invariant with respect to translations of the coordinate axes. In our investigation of b_L we shall make the assumption that the ordering according to the magnitude of the x_i is identical with the ordering according to the X_i . If that is so, the w_i are simply constants, and b_L is the ratio of two linear forms in x_i and y_i . In some fields of research such an assumption might be unrealistic, but in the physical sciences it would ordinarily be regarded as quite reasonable. In any event, occurence of situations where the assumption is reasonable seem sufficiently frequent to justify a detailed investigation.

We propose now to examine the bias and mean square error of b_L as they relate, in particular, to the choice of the w_i . We first set down the model and the assumptions which underlie the discussion in Sections 3, 4, and 5 below. We suppose that

$$Y_i = \alpha + \beta X_i , \qquad i = 1, 2, \cdots , n,$$

where α and β are unknown parameters while the X_i and Y_i are unknown constants. The X_i are assumed to be ordered according to increasing magnitude. We observe x_i and y_i , where

$$x_i = X_i + e_i ,$$

$$y_i = Y_i + f_i .$$

The e_i and f_i are random variables, representing errors of observation, such that

(1) e_i and e_j are independent if $i \neq j$,
(2) f_i and f_j are independent if $i \neq j$,
(3) e_i and f_j are independent for all i and j,
(4) $E(f_i) = 0$ and $E(f_i^2) = \sigma_f^2 = \nu_2$.
(5) Let $c = \min_i \mid X_{i+1} - X_i \mid$. Then Prob $\{\mid e_i \mid \geq \frac{1}{2}c\} = 0$. That is to say, the e_i have finite range, extending from $-\frac{1}{2}c$ to $\frac{1}{2}c$. This is the assumption which ensures that the ordering according to the x_i is identical with the ordering according to the X_i .
(6) $E(e_i) = 0$ and $E(e_i^2) = \sigma_e^2 = \mu_2$, $E(e_i^4) = \mu_4$, $E(e_i^6) = \mu_6$. All odd moments of the e_i are zero.
(7) $\sum w_i X_i \neq 0$. It is easily verified that for each of the estimators considered this is the case.

It should be stressed that the e_i and f_i represent errors of observation only since the X_i and Y_i are constants, otherwise e_i and f_i would obviously not be independent. As for assumption (5), it also ensures that the estimators are consistent, since $\sum w_i = 0$. In practice it is conceivable that one might wish to relax assumption (5) slightly to permit a distribution with infinite range but with most of the probability concentrated in $(-c/2, c/2)$. The extent to which one might relax this assumption is not considered in this paper but will be investigated subsequently.

3. BIAS AND MEAN SQUARE ERROR OF b_L.

The bias, B_L of b_L , may be obtained in the following way:

$$E(b_L) = E(\sum w_i y_i) E\left(\frac{1}{\sum w_i x_i}\right).$$

Now

$$\frac{1}{\sum w_i x_i} = \frac{1}{P + \sum w_i e_i}$$

$$= \frac{1}{P} \sum_{k=0}^{t} (-1)^k P^{-k} (\sum_i w_i e_i)^k + \frac{(-1)^{t+1}}{P^{t+2}} \frac{(\sum w_i e_i)^{t+1}}{1 + \dfrac{\sum w_i e_i}{P}},$$

where $P = \sum w_i X_i$.
Consequently,

$$E\left(\frac{1}{\sum w_i x_i}\right) = \frac{1}{P} \sum_{k=0}^{t} (-1)^k P^{-k} E(\sum w_i e_i)^k + D,$$

where

$$D = \frac{(-1)^{t+1}}{P^{t+2}} E\left\{(\sum w_i e_i)^{t+1} \Big/ \left[1 + \frac{\sum w_i e_i}{P}\right]\right\}.$$

It has proved sufficiently accurate for the purpose of this investigation to take $t = 5$. Then

$$E\left(\frac{1}{\sum w_i x_i}\right) = \frac{1}{P} + \frac{1}{P^3} E(\sum w_i e_i)^2 + \frac{1}{P^5} E(\sum w_i e_i)^4 + D,$$

where

$$D = \frac{1}{P^7} E\left[(\sum w_i e_i)^6 \Big/ \left(1 + \frac{\sum w_i e_i}{P}\right)\right].$$

Since $E(\sum w_i y_i) = \beta P$, we have

$$E(b_L) = \beta\left[1 + \frac{1}{P^2} E(\sum w_i e_i)^2 + \frac{1}{P^4} E(\sum w_i e_i)^4\right] + \beta P\, D.$$

Therefore

$$B_L = \beta\left[\frac{1}{P^2}\mu_2 \sum w_i^2 + \frac{1}{P^4}(\mu_4 \sum w_i^4 + 3\mu_2^2 \sum_{i \neq j}\sum w_i^2 w_j^2)\right] + \beta P\, D.$$

In all of the cases studied it is easily verified that

$$-1 < \sum w_i e_i / P < 1.$$

Consequently, for the cases we want to investigate it is possible to find simple bounds for D. In fact,

$$0 < D < \frac{1}{2P^7}\left[\frac{1}{1 - \frac{\frac{1}{2}\sum |\, w_i \,| c}{P}} + 1\right] E(\sum w_i e_i)^6.$$

To see this, one need only realize that for negative values of $\sum w_i e_i$

$$0 \leq \sum (w_i e_i)^6 \Big/ \left(1 + \frac{\sum w_i e_i}{P}\right) \leq \sum (w_i e_i)^6 \Big/ \left(1 - \frac{\frac{1}{2}c \sum |\, w_i \,|}{P}\right),$$

while for positive values of $\sum w_i e_i$

$$0 \leq \sum (w_i e_i)^6 \Big/ \left(1 + \frac{\sum w_i e_i}{P}\right) \leq \sum (w_i e_i)^6.$$

Consequently,

$$0 \leq E\left[\frac{\sum (w_i e_i)^6}{1 + \frac{\sum w_i e_i}{P}}\right] \leq \left[\frac{1}{1 - \frac{\frac{1}{2}c \sum |\, w_i \,|}{P}}\right] \tfrac{1}{2} E(\sum w_i e_i)^6 + \tfrac{1}{2} E(\sum w_i e_i)^6.$$

Therefore

$$0 < D < \frac{1}{2P^7}\left[\frac{1}{1 - \frac{\frac{1}{2}\sum |w_i| c}{P}} + 1\right] \cdot (\mu_6 \sum w_i^6 + 15\mu_4\mu_2 \sum_{i \neq j}\sum w_i^4 w_j^2 + 15\mu_2^3 \sum_{i \neq j \neq k}\sum \sum w_i^2 w_j^2 w_k^2).$$

Thus we have the means for approximating the bias to any described degree of accuracy, and furthermore, we have a way of assessing the accuracy of any approximation.

It is possible to obtain the mean square error, M_L of b_L, in essentially the same way as B_L was obtained. The result is

$$M_L = \beta^2\left[\frac{1}{P^2}\mu_2 \sum w_i^2 + \frac{3}{P^4}(\mu_4 \sum w_i^4 + 3\mu_2^2 \sum_{i \neq j}\sum w_i^2 w_j^2)\right] + \lambda \frac{1}{P^2}\left[\mu_2 \sum w_i^2 + \frac{3}{P^4}\mu_2^2(\sum w_i^2)^2 + \frac{5}{P^6}\sum w_i^2(\mu_4 \sum w_i^4 + 3\mu_2^2 \sum_{i \neq j}\sum w_i^2 w_j^2)\right] + \Delta, \quad \text{where} \quad \lambda = \nu_2/\mu_2$$

and

$$0 < \Delta < \left\{\beta^2 \frac{7}{2}\left(1 - \frac{\frac{1}{2}c \sum |w_i|}{P}\right)^{-2} + \frac{5}{2} - \left(1 + \frac{\frac{1}{2}c \sum |w_i|}{P}\right)^{-1} + \frac{\lambda\mu_2 \sum w_i^2}{P^2}\left[\frac{7}{2}\left(1 - \frac{\frac{1}{2}c \sum |w_i|}{P}\right)^{-2} + \frac{7}{2}\right]\right\} \cdot \{\mu_6 \sum w_i^6 + 15\mu_4\mu_2 \sum_{i \neq j}\sum w_i^4 w_j^2 + 15\mu_2^3 \sum_{i \neq j \neq k}\sum \sum w_i^2 w_j^2 w_k^2\}.$$

4. OPTIMAL CHOICE OF WEIGHTS.

We should like now to consider two possible sets of w_i; namely,

1. The set which minimizes the bias B_L, and
2. The set which minimizes the mean square error M_L.

In general, of course, the two sets differ. The exact specifications of either set does not appear possible in view of the complexity of the expressions for B_L and M_L. On the other hand it is possible to obtain very good approximate specifications.

If one chooses the w_i according to the scheme of Wald, or Bartlett, or Housner and Brennan, $\max_{e_i} |\sum w_i e_i|/P$ turns out to be consider-

ably less than unity, even for values of n as small as four; this is a consequence of assumption (5) of Section 2. This leads one to conjecture that $\max \mid \sum w_i e_i \mid / P$ is small for any reasonable choice of the w_i—in particular, one would expect it to be small for good choices, like that which minimizes the bias or the mean square error. If this is the case, one would expect the bias and the mean square error to be determined primarily by the term of order P^{-2} in the expression for each; consequently, minimizing that term should approximately minimize the bias or the mean square error, as the case may be. Now this term in the expression for the bias is $\mu_2 \beta P^{-2} \sum w_i^2$, whereas the corresponding term in the expression for the mean square error is $\mu_2(\beta^2 + \lambda)P^{-2} \sum w_i^2$. Obviously both of these terms are minimized by the same set of w_i, which is easily shown to be given by $w_i = (X_i - \bar{X})C$, where C is any arbitrary constant unequal to 0.

Up to this point we have proceeded heuristically; we should now like to show with the aid of some examples that retention of terms in P^{-4} makes only a trifling difference in the weights and far less difference in the bias itself.

As the first example, we consider four points: $X_1 = -3c/2$, $X_2 = -c/2$, $X_3 = c/2$, and $X_4 = 3c/2$. We take $\mu_2 = \frac{1}{3}(c/2)^2$ and $\mu_4 = \frac{1}{5}(c/2)^4$ which are the moments of the rectangular distribution on the interval $(-c/2, c/2)$, which is the maximum possible range consonant with the assumptions of Section 2. We might equally well consider any other admissible distribution on this finite range; the same technique would apply; and the results would be substantially the same. The condition above gives as the weights $w_1 = -3$, $w_2 = -1$, $w_3 = +1$, $w_4 = +3$ (or any multiple of these, of course). Minimization of

$$B_L \doteq \beta\left[\frac{1}{P^2}\mu_2 \sum w_i^2 + \frac{1}{P^4}(\mu_4 \sum w_i^4 + 3\mu_2^2 \sum \sum w_i^2 w_j^2)\right]$$

gives for the w_i

$$w_1 = -3.044, \qquad w_2 = -1, \qquad w_3 = +1, \qquad w_4 = +3.044.$$

This introduces a relative change in the bias $\Delta B_L / B_L$ of 9×10^{-5}.

As a second example, we consider the four points $X_1 = -5c/2$, $X_2 = -c/2$, $X_3 = c/2$, and $X_4 = 5c/2$, with μ_2 and μ_4 just as in the previous example. The condition gives

$$w_1 = -5, \qquad w_2 = -1, \qquad w_3 = +1, \qquad w_4 = +5.$$

Minimization of

$$B_L \doteq \beta\left[\frac{1}{P^2}\mu_2 \sum w_i^2 + \frac{1}{P^4}(\mu_4 \sum w_i^4 + 3\mu_2^2 \sum \sum w_i^2 w_j^2)\right]$$

gives for the w_i

$$w_1 = -5.034, \quad w_2 = -1, \quad w_3 = +1, \quad w_4 = 5.034.$$

This introduces a relative change in the bias $\Delta B_L/B_L = 3 \times 10^{-6}$.

It will be observed that the perturbations produced both in the weights and in the bias by the term in P^{-4} are smaller in the second example than in the first. That is exactly what one would expect, for although the spread of the X_i has increased in the second example, the range of the e_i has not changed. This has the effect of decreasing $\max \sum w_i e_i/P$, thereby decreasing the importance of the term in P^{-4} relative to that of the term in P^{-2}. One might examine the mean square error in similar fashion. It is clear that the approximate minimization of bias and mean square error given by choosing $w_i = X_i - \bar{X}$ is a very good approximation, and the estimator b_O which employs these weights can for all practical purposes, be regarded as having minimum bias and minimum means square error. We shall refer to this choice of weights as optimal and to the estimator

$$[\sum (X_i - \bar{X})y_i]/[\sum (X_i - \bar{X})x_i]$$

as the optimal (ratio-of-linear-forms) estimator.

5. COMPARISON OF VARIOUS WELL-KNOWN ESTIMATORS WITH THE OPTIMAL ESTIMATOR.

In general the spacing of the X_i is unknown, and the optimal estimator is therefore not obtainable. It is customary to use the weights of Wald, of Bartlett, or of Housner and Brennan. The corresponding estimators will henceforth be denoted by b_W , b_B , b_H respectively, and the corresponding biases and mean square errors will also carry these subscripts. For example B_B denotes the bias of Bartlett's estimator.

When the X_i are uniformly spaced, b_H and b_O are identical, which is to say that the Housner-Brennan estimator is optimal for this spacing. For any other spacing of the X_i neither b_H , b_B , nor b_W is optimal; the question is, how far do they depart from optimality.

To answer this question we have studied the bias and the mean square error of these estimators for $n = 6$, 12 and for various types of spacings of the X_i , which may be classified as follows:

1. Symmetric spacing: $\cdots$, $-5^{\omega}p$, $-3^{\omega}p$, $-1^{\omega}p$, $1^{\omega}p$, $3^{\omega}p$, $5^{\omega}p$, $\cdots$ for $\omega \geq 0$
2. Symmetric spacing: $-1^{\omega}p$, $-2^{\omega}p$, $-3^{\omega}p$, $\cdots$, $3^{\omega}p$, $2^{\omega}p$, $1^{\omega}p$ for $\omega < 0$
3. Asymmetric spacing: 0, $2(1^{\omega})p$, $2(2^{\omega})p$, $2(3^{\omega})p$, $\cdots$ for $\omega > 1$ and $0 < \omega < 1$
4. Asymmetric spacing: $-1^{\omega}p$, $-2^{\omega}p$, $-3^{\omega}p$, $\cdots$ for $\omega < 0$.

We have taken the values of p to make the ranges of the selected X_i values approximately the same, ensuring at the same time that condition (5) of Section 2 holds. In addition we have taken

$$\mu_2 = \frac{1}{3}\left(\frac{c}{2}\right)^2, \qquad \mu_4 = \frac{1}{5}\left(\frac{c}{2}\right)^4, \qquad \mu_6 = \frac{1}{7}\left(\frac{c}{2}\right)^6,$$

which coincide with the second, fourth, and sixth moments of a rectangular distribution having the maximum range consonant with the assumptions of Section 2. As indicated in Section 4 there is no serious loss of generality in confining our attention to a rectangular distribution of errors.

Tables 1 to 5 present, for $n = 6$ and $n = 12$,

1. The bias of the optimal estimator.
2. The ratio of the bias of b_W, b_B, and b_H to the bias of the optimal estimator.
3. The ratio of the mean square error of b_W, b_B, and b_H to that of the optimal estimator.

The values cited in the tables are correct to within one unit in the last place given there. Results for other values of n can be found in Dorff [2].

6. LEAST-SQUARES-TYPE ESTIMATOR

This section deals with $b_Q = [\sum(x_i - \bar{x})y_i]/[\sum (x_i - \bar{x})x_i$, which is the estimator one would obtain if he minimized the sum of squares of deviations in the vertical direction, simply ignoring the fact that the x_i are random variables. It is well-known that b_Q is not a consistent

TABLE 1

MAXIMUM VALUE OF B_O/β.

n	Values of ω 2.0	1.0	0.5	−0.5	−2.0
			Symmetric Spacing		
6	.0002	.0048	.0012	.0004	.0008
12	.0000	.0006	.0001	.0001	.0000
			Asymmetric Spacing		
6	.0002	—[3]	.0014	.0006	.0000
12	.0000	—	.0002	.0000	.0000

[3]For $\omega = 1.0$ the pattern of spacing presented would be symmetric; consequently no value is cited here.

TABLE 2

BIAS RATIO FOR SYMMETRIC SPACING

Ratio	Values of ω 2.0	1.0	0.5	−0.5	−2.0
	$n = 6$				
B_H/B_O	1.06	1.00	1.04	1.08	1.10
B_W/B_O	1.73	1.30	1.09	1.06	1.74
B_B/B_O	1.22	1.10	1.14	1.25	1.38
B_Q/B_O	−3.00	−2.97	−2.90	−2.99	−2.99
	$n = 12$				
B_H/B_O	1.06	1.00	1.04	1.06	1.47
B_W/B_O	1.78	1.32	1.11	1.11	2.93
B_B/B_O	1.28	1.12	1.15	1.26	2.13
B_Q/B_O	−9.00	−9.00	−9.00	−9.00	−9.00

TABLE 3

BIAS RATIO FOR ASYMMETRIC SPACING

Ratio	Values of ω 2.0	0.5	−0.5	−2.0
	$n = 6$			
B_H/B_O	1.08	1.10	1.17	1.62
B_W/B_O	1.41	1.58	1.69	2.81
B_B/B_O	1.19	1.27	1.34	2.03
B_Q/B_O	−3.00	−3.00	−3.00	−3.00
	$n = 12$			
B_H/B_O	1.08	1.09	1.32	2.51
B_W/B_O	1.43	1.56	2.08	5.24
B_B/B_O	1.20	1.27	1.61	3.66
B_Q/B_O	−9.00	−9.00	−9.00	−9.00

TABLE 4

MEAN-SQUARE-ERROR RATIO FOR SYMMETRIC SPACING

Ratio	Values of ω 2.0	1.0	0.5	−0.5	−2.0
	$n = 6$				
M_H/M_O	1.06	1.00	1.04	1.08	1.10
M_W/M_O	1.73	1.31	1.09	1.06	1.74
M_B/M_O	1.22	1.10	1.14	1.26	1.38
M_Q/M_O	0.99	—[1]	0.99	0.99	1.00
	$n = 12$				
M_H/M_O	1.06	1.00	1.04	1.07	1.47
M_W/M_O	1.78	1.32	1.11	1.11	2.93
M_B/M_O	1.28	1.12	1.15	1.26	2.13
M_Q/M_O	1.00	1.01	1.00	1.00	1.00

[1]Because of slow convergence of our approximation techniques for this combination of parameters, we have no reliable value to cite here.

TABLE 5

MEAN-SQUARE-ERROR FOR ASYMMETRIC SPACING

Ratio	Values of ω 2.0	0.5	−0.5	−2.0
	$n = 6$			
M_H/M_O	1.09	1.11	1.17	1.62
M_W/M_O	1.41	1.58	1.69	2.81
M_B/M_O	1.19	1.27	1.34	2.03
M_Q/M_O	1.00	0.98	1.00	1.00
	$n = 12$			
M_H/M_O	1.08	1.09	1.32	2.51
M_W/M_O	1.43	1.56	2.08	5.24
M_B/M_O	1.20	1.27	1.61	3.66
M_Q/M_O	1.00	1.00	1.00	1.00

estimator of β; however, it is conceivable that it might have desirable small-sample properties. We shall now investigate the bias and mean square error of b_Q , making the same assumptions as in Section 2. Tables 2–5 cite values of B_Q/B_O and M_Q/M_O correct to within one unit in the last place given there. It will be observed that B_Q/B_O is approximately $-(n-3)$ while M_Q/M_O is nearly unity.

7. DISCUSSION OF ESTIMATORS WHICH DO NOT REQUIRE REPLICATION

We should like now to compare the estimators discussed in Section 2 with a view toward making some practical suggestions for workers who are actually concerned with fitting straight-line relations.

Generally speaking b_B is clearly superior to b_W but there are some situations when such is not the case; we found that b_W had lower bias and lower mean square error than b_B when the X_i are symmetrically spaced, with $\omega = 0.5$ or $\omega = -0.5$. However, in almost every case investigated, b_H proved superior to b_W and to b_B . This, of course, is hardly surprising. The use of b_H presupposes through the ability to order all the observations more information available to the experimenter than that of b_W or b_B , and estimators based upon more complete information are typically more efficient than those based upon partial information. We did, nevertheless, find instances where b_W proved superior to b_B and b_H . One of these is $n = 6$, $\omega = -0.5$, symmetric spacing. For asymmetric spacings corresponding to these values of n and ω, b_H was found decidedly superior to the others.

As the comparison of b_H and b_Q is slightly more involved, we should like to first compare b_O with b_Q . The conclusion seems inescapable that when b_O can be obtained it is preferable to b_Q for two reasons:

1) b_O has much smaller bias than b_Q , while its mean square error is essentially no greater.
2) b_O is a consistent estimator of β, whereas b_Q is not.

The difficulty, of course, is that b_O cannot usually be obtained, inasmuch as the spacing of the X_i is unknown.

However it appears that b_H is a reasonable substitute for b_O whenever the X_i possess no marked skewness and are not bunched excessively. When such is the case we see that neither B_H/B_O nor M_H/M_Q will greatly exceed unity, while B_H/B_Q will be approximately $-1/(n-3)$. Thus, when using b_H in preference to b_Q , one stands to do much better with respect to bias and very little worse with respect to mean square error.

One could roughly assess the skewness or bunchiness of his data by

plotting the x_i versus i on log-log paper; the slope of the line which best fits the points is then ω. Ordinarily this should not be necessary; in most cases one could decide whether there was evidence of skewness or bunchiness by merely looking at the data.

8. ESTIMATORS WHICH REQUIRE REPLICATION.

We consider now various estimators of β suitable when replicated observations are available. We suppose that

$$Y_i = \alpha + \beta X_i , \qquad i = 1, 2, \cdots , n,$$

just as in Section 2, but now we have observations

$$x_{it} = X_i + e_{it} ,$$

$$y_{it} = Y_i + f_{it} ,$$

where $t = 1, 2, \cdots , r (r > 1)$;
e_{it} and f_{it} are random variables, representing errors of observation, such that

1) e_{it} and e_{ju} are independent unless $i = j$ and $t = u$.
2) f_{it} and f_{ju} are independent unless $i = j$ and $t = u$.
3) e_{it} and f_{ju} are independent for all i, j, t, u.
4) Let $c = \min_i | X_{i+1} - X_i |$. Then Prob $\{| e_{it} | \geq c/2\} = 0$.
5) $E(e_{it}^2) = \mu_2$, $E(e_{it}^4) = \mu_4$, $E(e_{it}^6) = \mu_6$. All odd moments of the e_{it} are zero.
6) Prob $\{| f_{it} | \geq \frac{1}{2}h\beta c\} = 0$. That is to say, we now restrict f_{it} to a finite range $(-\frac{1}{2}h\beta c, \frac{1}{2}h\beta c)$; the symbol h is defined by this relation.
7) $E(f_{it}^2) = \nu_2 = h^2\beta^2\mu_2$, $E(f_{it}^4) = \nu_4 = h^4\beta^4\mu_4$, $E(f_{it}^6) = \nu_6 = h^6\beta^6\mu_6$. All odd moments of the f_{it} are zero. (These assumptions would be satisfied if the distribution of the f_{it} differed from that of the e_{it} only in its range.)

In our previous paper [3] we considered the following estimators of β, all of which are consistent provided that their denominators have non-vanishing expectations:

$$b_1 = B_{xy}/(B_{xx} - W_{xx}), \quad b_2 = (B_{yy} - W_{yy})/B_{xy} , \quad b_3 = \sqrt{\frac{B_{yy} - W_{yy}}{B_{xx} - W_{xx}}} \operatorname{sgn} \beta$$

where

$$B_{xx} = \frac{r}{n-1} \sum (x_{i\cdot} - x_{\cdot\cdot})^2, \qquad W_{xx} = \frac{1}{n(r-1)} \sum \sum (x_{it} - x_{i\cdot})^2,$$

$$B_{yy} = \frac{r}{n-1}\sum (y_{i.} - y_{..})^2, \qquad W_{yy} = \frac{1}{n(r-1)}\sum\sum (y_{ij} - y_{i.})^2,$$

$$B_{xy} = \frac{r}{n-1}\sum (x_{i.} - x_{..})(y_{i.} - y_{..}),$$

and $\operatorname{sgn}\beta = +1$ if $\beta > 0$
-1 if $\beta < 0$.

We have applied the techniques of Section 3 to the study of the bias and mean square error of these three estimators for comparison with the bias and mean square error of $b_O = [\sum (X_i - \bar{X})y_i]/[\sum (X_i - \bar{X})x_i]$. We found that b_1 , b_2 , b_3 and b_O have approximately the same mean square error but that their biases can differ greatly. In fact, when the errors are sufficiently small $B_1/B_O \doteq 2$, $B_2/B_O \doteq 1 - h^2$, $B_3/B_O \doteq (3 - h^2)/2$. Thus, the absolute bias of b_1 is smaller than that of b_2 whenever $h > \sqrt{3}$; the absolute bias of b_3 is smaller than that of b_1 and b_2 whenever $\sqrt{5/3} < h < \sqrt{7}$. Now $h^2 = \nu_2/\mu_2/\beta^2$, and since ν_2/μ_2 is ordinarily unknown, we are unable to decide whether $\sqrt{5/3} < h < \sqrt{7}$ or not. This would seem to rule out b_3 as a useful estimator of β. On the other hand in many problems it might well be possible to decide whether $h > \sqrt{3}$ or not. Inasmuch as b_1 and b_2 have approximately the same mean square error one could then use the bias as a criterion for choosing between the two: i.e.,

$$\text{When } \beta^2 < \frac{\nu_2/\mu_2}{3}, \quad \text{use } b_1 ;$$

$$\text{When } \beta^2 > \frac{\nu_2/\mu_2}{3} \quad \text{use } b_2 .$$

It is interesting to note that analysis of the asymptotic variance of b_1 and b_2 in our previous paper [3] led us to recommend that

$$\text{when } \beta^2 < \nu_2/\mu_2 , \quad \text{use } b_1 ,$$

$$\text{when } \beta^2 > \nu_2/\mu_2 , \quad \text{use } b_2 ,$$

so that the large sample approach based on variance and the small sample approach based on bias lead to somewhat different conclusions.

For values of h larger than $\sqrt{2}$, b_O would be preferable to b_1 , but unfortunately b_O is not usually obtainable. We could, of course, use b_H , b_W , or b_B in lieu of b_1 at some sacrifice of efficiency, but with a corresponding reduction of the magnitude of the bias. If one knew

something about the spacing of the X_i, this might be a desirable way to obtain an estimate of β, but if nothing is known about the X_i, it seems advisable to use either b_1 or b_2. In any case the use of b_1 or b_2 does not require any assumption about the possibility of ordering the observed values correctly in contradistinction to the estimators b_H, b_W, b_B, or b_O.

9. CORRELATED ERRORS

In some investigations it would not be reasonable to assume the errors e_i and f_i to be independent; one should properly take account of correlation in such cases. The correlation of errors greatly increases the complexity of the problem and an examination of small sample properties for such situations has not been included here.

We have, however, examined in [2], [3] the asymptotic variance for consistent estimators of β when the errors are correlated. In this case, the estimators b_1, b_2, and b_3 are modified as follows in order to retain the property of consistency:

$$b_1 = (B_{xy} - W_{xy})/(B_{xx} - W_{xx}), \qquad b_2 = (B_{yy} - W_{yy})/(B_{xy} - W_{xy}),$$

$$b_3 = \sqrt{(B_{yy} - W_{yy})/(B_{xx} - W_{xx})}\ \operatorname{sgn}\beta.$$

Essentially the same conclusions were obtained as in the case of independent errors:

1. When $\beta^2 < \nu_2/\mu_2$, use b_1,
2. When $\beta^2 > \nu_2/\mu_2$, use b_2,
3. b_3 is advantageous for only a narrow range of values of the parameters and is therefore not likely to be of much interest.

10. CONCLUSIONS

In this paper we have considered the problem of estimating the slope of a linear functional relation when both variables are in error.

When replicated observations are made a number of estimators are available, all of which have approximately the same asymptotic variance and mean square error, but study of the bias enables us to choose among them when we have partial information concerning the parameters.

When there is no replication of observations and the errors are suitably restricted, there exists a class of consistent estimators which are ratios of linear forms in the observations. When partial information is available concerning the spacing of the X_i there are rational grounds for choosing a suitable member of the class.

BIBLIOGRAPHY

[1] Bartlett, M. S. [1949]. Fitting a straight line when both variables are subject to error. *Biometrics* 5.

[2] Dorff, M. R. [1960]. Large and small sample properties of estimators for a linear functional relation. *Unpublished Ph. D. Thesis*, Iowa State University of Science and Technology.

[3] Dorff, M. R. and Gurland, J. [1961]. Estimation of the parameters of a linear functional relation, Accepted for publication in Journal of the Royal Statistical Society, Series B.

[4] Gibson, W. M. and Jowett, G. H. [1957]. 'Three-Group' regression analysis. Part I. Simple Regression Analysis. *Appl. Statist. 6.*

[5] Gurland, J. [1956]. *Paper Presented at meeting of Institute of Mathematical Statistics Chicago.*

[6] Hooper, J. W. and Theil, H. [1958]. The extension of Wald's method of fitting straight lines to multiple regression. *Rev. Int. Statist. Inst. 26.*

[7] Housner, G. W. and Brennan, J. F. [1948]. Estimation of linear trends. *Ann. Math. Statist. 19.*

[8] Nair, K. R. and Shrivastava, M. P. [1942]. On a simple method of curve fitting. *Sankhya 6.*

[9] Theil, H. and van Yzeren, J. [1956]. On the efficiency of Wald's method of fitting straight lines. *Rev. Int. Statist. Inst. 24.*

[10] Wald, A. [1940]. The fitting of straight lines if both variables are subject to error. *Ann. Math. Statist. 11.*

STATISTICS FOR A DIAGNOSTIC MODEL[1]

ADRIANUS J. VAN WOERKOM AND KEEVE BRODMAN
Cornell University Medical Center
New York, New York, U. S. A.

INTRODUCTION

In recent years, several methods have been proposed for making medical diagnoses by machine (Ledley and Lusted [1959], Crumb and Rupe [1959]). A method devised by Brodman *et al.* [1959, 1960] has been used to program a high-speed electronic computer for making presumptive medical diagnoses using only information relating to the age, sex, and responses of patients to a standardized health questionnaire.

The method assigns patients to none, one, or several of 60 selected disease categories. The 60 diseases most frequently diagnosed by hospital physicians in men and the 60 diseases in women were chosen for study. The method was developed with data referring to 5,929 consecutive adult white patients (2,718 men and 3,211 women) admitted to the outpatient departments of The New York Hospital, a large general hospital, during the 18-month period beginning July 1, 1948. It was tested with data referring to 2,745 consecutive adult white patients (1,280 men and 1,465 women) admitted during the 12-month period beginning January 1, 1956.

Each patient's symptoms were elicited through a printed form, the Cornell Medical Index-Health Questionnaire (CMI). The CMI was devised to collect diagnostically important elements of the medical history given by general medical patients, without expenditure of a physician's time. Solely with these data, a physician can often correctly predict which diseases will be found in subsequent examination (Brodman *et al.* [1951]).

Additional data abstracted for analysis from the hospital records include each patient's sex and age, along with the diagnoses made by hospital physicians after eliciting a history and performing physical and laboratory examinations. These diagnoses for the 1948–1949 data were coded according to the U. S. Public Health Service *Manual for Coding Causes of Illness* [1944] and form the standard against which the method is evaluated.

[1]Leo S. Goldstein and Albert J. Erdmann Jr. participated in this study, which was carried out at the Department of Medicine, New York Hospital-Cornell Medical Center with the support of a Public Health Service Research Grant.

The problem may be phrased in the following terms: Patients in the sample may belong to a number of specified disease categories. Patients were assigned by physicians to disease categories with a low incidence of false positives (incorrect diagnoses) but were not assigned to every category to which they belonged. Each patient has a number of attributes (complaints on the CMI). These attributes may or may not be correlated with each other. No one-to-one relation exists between categories and attributes. The problem is to assign patients to categories solely by means of attributes indicated on the CMI, with a maximum of correct and a minimum of incorrect assignments.

There are two obvious approaches to the solution of this problem:

1. The use of the conditional probabilities of a patient's having an attribute, given the category to which he belongs. From these, using Bayes' theorem, the conditional probability can be found for the patient's belonging to a specific category, given the attributes (Feller [1957], Chapter 5).

2. The use of significance values or weights for the attributes with respect to the categories.

CONDITIONAL PROBABILITIES

Let P_k be the probability or relative frequency of attribute A_k in the total sample, and p_{ik} the conditional probability of the attribute A_k given the category D_i. Then

$$p_{ik} = \Pr(A_k \mid D_i). \tag{1}$$

If n_i is the sub-sample size of category D_i and N the size of the total sample, then

$$\Pr(D_i) = n_i/N. \tag{2}$$

The conditional probability for a group of attributes $A_{k_1}, \cdots, A_{k_n}$ is:

$$\Pr(A_{k_1} \cdots A_{k_n} \mid D_i) = \Pr(A_{k_1} \mid D_i) \Pr(A_{k_2} \mid D_i) \cdots \Pr(A_{k_n} \mid D_i). \tag{3}$$

Using Bayes' theorem, the probability of the patient's belonging to category D_i given the attribute complex $A_{k_1}, \cdots, A_{k_n}$ is:

$$\Pr(D_i \mid A_{k_1} \cdots A_{k_n}) = \frac{\Pr(A_{k_1} \mid D_i) \cdots \Pr(A_{k_n} \mid D_i) \Pr(D_i)}{\sum_i \Pr(A_{k_1} \mid D_i) \cdots \Pr(A_{k_n} \mid D_i) \Pr(D_i)}. \tag{4}$$

The expression (4) for Bayes' rule will have to be adjusted for correlation between the attributes. As can be seen from Table 1, the sample sizes of the 60 selected diseases are generally quite small. These sample sizes are of the order of 30, which makes the measurements of the

TABLE 1

SAMPLE SIZES OF 60 DISEASES IN MEN AND WOMEN.

Sample	Men	Women
Size	N	N'
10–19	16	8
20–29	13	12
30–39	8	14
40–49	6	6
50–99	10	12
100–149	6	5
150–404	1	3

correlation coefficients uncertain. In Table 2 a tabulation is given for the correlation coefficients for the symptoms with higher proportionality in the disease sub-samples than in the total sample. The coefficients are reduced to unit variance. Only 22 percent of these correlations are more than twice their standard error.

In general, a patient has more than two symptoms and multiple correlations will have to be considered. Because of the small sub-sample sizes, these multiple correlations, either directly observed or

TABLE 2

DISTRIBUTION OF ITEM INTER-CORRELATIONS REDUCED TO UNIT VARIANCE.

Correlation/Variance	N
4.0–6.0	12
3.0–4.0	20
2.0–3.0	43
1.5–2.0	43
1.0–1.5	41
0.5–1.0	57
0.0–0.5	38
0.0–(−0.5)	44
(−0.5)–(−1.0)	24
(−1.0)–(−1.5)	13
(−1.5)–(−2.0)	3
(−2.0)–(−3.0)	0
(−3.0)–(−4.0)	0
(−4.0)–(−6.0)	0

TABLE 3

NUMBER OF DISEASES DIAGNOSED PER PATIENT (MAXIMUM OF 3).

Number of Diseases	*N*		%	
	Men	Women	Men	Women
1	1456	1635	54	51
2	646	811	24	25
3	616	765	22	24

derived from the covariances, will be more uncertain than the zero order correlations and their use could lead to a large amount of bias. In view of this uncertainty and the tremendous amount of work involved, the measurement of all the correlations needed for the proper application of Bayes' rule is not justified.

Of particular interest with respect to the use of Bayes' rule, many patients have more than one disease. In Table 3 a distribution of diagnosed diseases is given for the 1948–1949 New York Hospital population.

The presence in a patient of more than one disorder considerably complicates a method using conditional probabilities. Symptoms claimed by a patient with a particular diagnosed disease may be related to other and often undiagnosed diseases. Due to a random distribution of most diseases with respect to each other, the co-existence of diseases in a patient will not seriously obscure the cluster of symptoms having substantially higher proportionality within a disease category than in the total sample, nor will the co-existence of diseases seriously distort the correlations of these symptoms within the disease category. Correlations between symptoms which have an incidence within the disease category similar to the incidence in the total sample, however, may be due to other diseases.

Some of these difficulties could be avoided by studying disease complexes rather than single diseases. It is evident from the sample sizes of single diseases as shown in Table 1, however, that the study of disease complexes in the sample available would not have been possible because of the small sub-sample sizes for these complexes. Even in a sample much larger than the one available, compounding of error because of false negatives would still be uncontrolled.

The effect on the conditional probabilities of the presence of diseases other than the one being studied would be partially eliminated by using only those attributes which have a substantially higher incidence within

TABLE 4

HYPOTHETICAL SAMPLE FOR CONDITIONAL PROBABILITIES.

	D_1	D_2
n	90	10
A_1	33%	60%
A_2	33%	60%
$n(A_1A_2)$	9	4
$\Pr(D\|A_1A_2)$	9/13	4/13

the disease category than in the total sample. This approach would also eliminate those symptoms which occur less frequently in the disease than in the total sample. (In clinical practice, these symptoms are given little diagnostic weight.)

Another objection to the use of Bayes' rule is that the conditional probabilities for a disease, given a set of attributes, is directly proportional to the sub-sample size of the disease category. (The medical profession first considers diseases with a high incidence but does not make a diagnosis on this basis.)

If the data available are consistent and complete, Bayes' rule will give a mininum number of wrong assignments. If the data are inconsistent and incomplete, as in the sample studied, there may be many cases of wrong assignments and methods other than Bayes' might be more effective. Referring to Table 4, if the 9 cases in D_1 with attributes A_1 and A_2 were false negatives in the data with regard to D_2 , the hypothesis that all patients having A_1 and A_2 belong to D_2 will be more effective than the use of Bayes' rule. This hypothesis can be tested only by re-examining hospital records or re-calling patients.

Notwithstanding the objections raised, the conditional probability method, with the modifications described above, was tested with the 1948–1949 sample of male patients. The results showed a number of expected inconsistencies. Nearly always, those categories with the largest number of attributes with high relative frequencies were favored. In addition, categories with small sample sizes and with a small number of attributes of high relative frequency (that would have been identified by a physician in many cases) failed to get probabilities high enough to be considered.

The most reliable information in the data is the diagnoses made by physicians after they have examined the patients. The one thing that can be established with minimum error from the data studied here is the significance of an attribute with respect to a disease. The con-

ditional probability approach fails to make use of this information.

It was found from clinical experience and from noting the relative frequencies or proportionalities P_k and p_{ik} that a disease category is characterized by a cluster of attributes each with a much higher proportionality in the disease category than in the total sample. If a patient's cluster of complaints is similar to the cluster of complaints of the average patient in that disease category, he can be assigned to the category. This approach is discussed in the following section.

SIGNIFICANCE VALUES

Let the variable S'_{ik} be given as

$$S'_{ik} = n_i(p_{ik} - P_k)/\sqrt{n_iP_k(1 - P_k)}, \tag{5}$$

where the notations are as before. It can be shown that in a sample where the attribute A_k is randomly distributed with probability P_k, the variable S'_{ik} is approximately normally distributed if $n_iP_k(1 - P_k)$ is not too small. However, according to Feller [1957 p. 170], the error committed in replacing the binomial distribution by the normal is surprisingly small, even for values of $n_iP_k(1 - P_k)$ of the order of 1.5. The comparison between the two for the probabilities of the variable between two boundaries is by far not as good, but this is of little consequence in this case. Then in the sub-sample of size n_i, the probability of observing the value S'_{ik} by chance is proportional to the expression $\exp(-S'^2_{ik})$. The observed value of S'_{ik} in the sub-sample of category D_i therefore expresses a measure of significance in the usual sense. Because the constant n_i occurs in the significance values for all attributes with regard to the category D_i, this factor can be dropped from the expression. Then,

$$S''_{ik} = (p_{ik} - P_k)/\sqrt{P_k(1 - P_k)} \sim (p_{ik} - P_k)/\sqrt{P_k}\,. \tag{6}$$

The term $(1 - P_k)$ varies between .99 and .46 with nearly all values around .80 with the square root equal to .9. By ignoring the term $(1 - P_k)$ an error is made of approximately 10 percent. This error is of little consequence because of statistical inaccuracies in the other parameters.

Let all attributes A_k for which $S''_{ik} \geq 2$ be called significant attributes. If the complex of significant attributes contains more than one attribute, $\sum S''^2_{ik}$ represents a measure of the joint probability of the attribute complex and thus the significance of the complex. The significance value of the complex should also include the correlation between the attributes. As mentioned before, however, in the sample under

study these correlations are generally insignificant and a linear model can be used.

The variable S''_{ik} measures not only the significance of the attributes but also their correlation with respect to the categories. The observed coincidence of category D_i and attribute A_k is equal to $(n_i/N)p_{ik}$ and the product of the expected values is $(n_i/N)P_k$. The co-variance is then equal to

$$(n_i/N)p_{ik} - (n_i/N)P_k = n_i/N(p_{ik} - P_k). \tag{7}$$

In order to reduce (7) to the correlation coefficient, the expression has to be divided by the square root of the product of the variances of n_i/N and P_k. The factor n_i/N is constant for all attributes in the same category and therefore can be eliminated. The variance of P_k depends on the universe from which our sample is drawn. If the universe has a Poisson distribution the value is equal to P_k; if the universe is binomial the value is equal to $P_k(1 - P_k)$.

The resulting expressions

$$(p_{ik} - P_k)/\sqrt{P_k} \quad \text{or} \quad (p_{ik} - P_k)/\sqrt{P_k(1 - P_k)}$$

are identical to S''_{ik} (6).

The following form was adopted for S_{ik}:

$$S_{ik} = [(p_{ik} - P_k)/2\sqrt{P_k}] - 1, \tag{8}$$

where the factor 2 in the denominator yields a convenient scale. A minimum value of 3 was originally set for $\sqrt{P_k}$. Further study of the data indicates that setting a minimum value for this is unnecessary, and that all negative values can be set to zero.

If the attributes $A_{k_1}, \cdots, A_{k_n}$ are present in a patient, the corresponding sums

$$(Sc)_i = \sum_{k_1}^{k_n} S^2_{ik} \qquad (j = 1, \cdots, 60) \tag{9}$$

are called the scores. The scores for a patient are normalized by dividing them by the mean score for the entire category. This normalization compares the patient to the average patient. If the sample sizes had been kept in the variables S_{ik} they would have cancelled out at this point.

It cannot be assumed that the score as indicated in formula (9) will give optimal results. The score is essentially a measure of the probability of the chance existence of any sub-sample with the same relative frequency p_{ik} for the attributes $A_{k_1}, \cdots, A_{k_n}$ as in the sub-sample of category D_i. When the score is high this probability is low,

that is, a high score indicates a small probability for the patient's not belonging to the category. Had the linear sum of the significance values S_{ik} been used to determine the score, small changes in the significance values would be reflected by large changes in the probabilities. This would make the score a bad indicator of the probabilities. If powers of S_{ik} higher than the second were used, the score would be a bad indicator for the opposite reason.

Various models (linear sum, sum of squares, sum of higher powers, etc.) may be compared quantitatively only if the distribution of scores in both the restricted and the category sub-samples are known. By restricted sample is meant the total sample minus the category sub-sample. These distributions could not be obtained for the sample studied, because the restricted sample contains patients who really belong to the category sub-sample but have not been assigned to it in the hospital (false negatives). Some of the disease categories, further, are significantly correlated. These correlated category sub-samples must be removed from the restricted sample in order to get the desired distribution. Removal of false negatives and adjunct category sub-samples is impossible for a large sample.

Physicians, when making a diagnosis, consider the patient's age. Table 5 shows the age distribution by decades of the 15 diseases diagnosed most frequently in men and in women in the 1948–1949 population. The relative frequency of a disease category is a function of age. The use of these relative frequencies in the sense of relative probabilities is not consistent with a model based on significance values of the attributes. There is, however, another aspect to be considered. Let the sample be subdivided into a number of age groups. The relative sub-sample sizes $(r_i)_m$ for disease category D_i and age group m will in general not be constant with respect to m. The significance value of an attribute with respect to a category is a function of the sub-sample size. If no subdivision of the total sample is considered, the effect of the sub-sample size is eliminated because the score of the patient is always compared to the score of the average patient and the sub-sample size does not have to enter into the significance value. If the same comparison is made in case a subdivision exists, the sub-sample size will have to be considered in the significance value. The expression for $(S_{ik})_m$ will be

$$(S_{ik})_m = \frac{\sqrt{(r_i)_m}(p_{ik} - P_k)}{\sqrt{P_k(1 - P_k)}} \sim \frac{\sqrt{(r_i)_m}(p_{ik} - P_k)}{\sqrt{P_k}}. \qquad (10)$$

The mean scores for each age group within the disease category were computed to test the possibility that the relative frequency p_{ik}

of the attributes within the category sub-sample is a function of age. No significant interaction was found between the mean scores and age.

In order to smooth the relative frequencies of the disease categories with respect to age, moving averages of three successive decades were used. After the $\sum S_{ik}^2$ scores were formed, using formula (8) for S_{ik}, the scores were multiplied by $(r_j)_m$.

This modification of the significance value changes the distributions of the scores for both the categories and total sample so that more success can be expected from the model. Indications of this are made in Figure 1, where *I* is the distribution curve for the restricted sample and *II* is the distribution curve for the sub-sample category. The dotted lines show the distribution curves using the formula modified for age. The trend will be primarily towards a smaller number of wrong assignments (false positives).

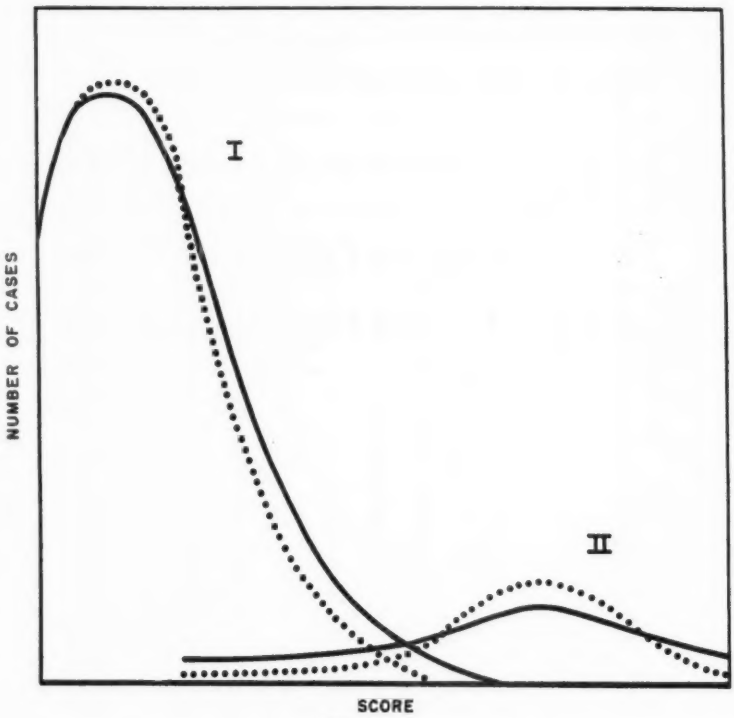

FIGURE 1

DISTRIBUTION OF SCORES FOR CATEGORY AND TOTAL SAMPLE.

TABLE 5

AGE DISTRIBUTION OF PATIENTS WITH SELECTED DIAGNOSES.

MEN

Diagnosis	Code	*N*	Percent of *N* by Decade						
			10–9	20–9	30–9	40–9	50–9	60–9	70+
Total		2718	05	16	18	20	22	14	04
Psychoneurosis	330	233	04	26	24	26	15	04	00
Benign prostatic hypertrophy	630	140	00	00	02	07	34	37	20
Inguinal hernia	550	120	04	03	11	28	24	23	07
Coronary artery disease	382	110	00	00	01	19	25	37	18
Varicose veins	410	110	01	05	15	22	29	23	05
Duodenal ulcer	527	109	02	11	27	30	18	09	03
Sinusitis	495	100	12	21	21	19	22	04	01
Functional gastro-intestinal disorder	564	96	00	25	22	28	16	09	00
Hemorrhoids	415	83	04	11	24	23	25	12	01
Errors of refraction	343	80	15	26	16	25	10	06	01
Infected teeth	510	80	11	14	25	20	19	09	03
Osteo-arthritis	724	76	00	00	03	17	39	32	09
Hypertensive cardiovascular disease	370	72	00	00	07	21	40	26	06
Arteriosclerosis	400	65	00	00	00	08	29	46	17
Deflected nasal septum	496	62	10	32	24	18	15	02	00

TABLE 5 (*Continued*)

WOMEN

Diagnosis	Code	*N*	Percent of *N* by Decade						
			10–9	20–9	30–9	40–9	50–9	60–9	70+
Total		3211	07	19	25	23	11	12	02
Psychoneurosis	330	404	06	23	35	22	08	06	00
Obesity	241	202	08	13	20	26	16	15	00
Varicose veins	410	168	00	05	18	31	18	24	03
Functional gastro-intestinal disorder	564	112	02	19	35	24	13	06	02
Errors of refraction	343	111	17	23	14	25	10	12	00
Hypertensive cardiovascular disease	370	111	00	00	05	21	33	29	12
Cervicitis	652	110	02	22	38	34	05	00	00
Dermatitis venenata	718	103	09	15	19	20	19	13	05
Sinusitis	495	96	06	28	31	23	07	04	00
Osteo-arthritis	724	96	00	00	04	28	23	33	11
Hemorrhoids	415	74	00	18	30	32	07	11	03
Infected teeth	510	71	18	18	30	20	06	08	00
Biliary calculi	585	68	00	10	24	29	15	22	00
Menstrual disorders	664	67	16	33	25	25	00	00	00
Allergic rhinitis (hay fever)	500	64	17	44	20	13	03	02	02

Because the process of assignment by physicians at The New York Hospital for the sample of patients under study cannot have been entirely divorced from the attributes, bias in assignment may be expected with respect to the attributes. Bias in p_{ik} arises partly as a function of the clinical process used in making a decision. A physician in oral interview elicits a patient's symptoms and then investigates with physical and laboratory examinations the diseases he presumes to be associated with the symptoms. Some of these symptoms will be similar to those indicated by the patient on the CMI. The presumptive diagnoses suggested by the symptoms reported to the physician can thus be expected to be related to the diagnostic assignments made from the attributes on the CMI. Bias in p_{ik} arising from the clinical process is inherent in the data and cannot easily be corrected.

In Table 6 is given a sample of the relative frequencies for 27 selected items on the CMI with respect to 6 selected diseases for men in the 1948–1949 population. The numbers in the heading of the table correspond to the item numbers on the questionnaire. The relative frequencies are expressed in percent. In the 7th line of the table the corresponding relative frequencies are given for the entire 1948–1949 sample of male patients. The last six lines of the table give the corresponding significance indices (S_{ik}). As can be seen, the complaints for the 6 selected diseases have very little overlap and form a good basis for differentiation of the diseases. The disease categories and items in the table were selected to illustrate this characteristic. The table shows that differential classifications can be made on the basis of the significance indices.

The sample in Table 7 was chosen from the male 1948–1949 population to present a case where overlap among diseases exists in the complaints. In the first column of the table the item content is given with the item number from the questionnaire. Column headings identify the diseases. The last column gives the number of diseases for which the particular item is significant and the last line in the table shows the number of these items significant for the disease. All significance indices smaller than 3 are omitted from the table. Even with the overlap present, the table still shows that differentiation among diseases can be made on the basis of the significance indices. Table 8 gives a comparable illustration for the women in the 1948–1949 population.

Some points in Tables 7 and 8 are of interest:

1. In Table 7, item 56 of disease 121 ($N = 12$) illustrates bias introduced because of small sample size in the disease.

2. In Table 8, item 41 of disease 382 is related to age.

TABLE 6

PERCENT "YES" AND SIGNIFICANCE INDICES FOR ITEMS WITH LITTLE OVERLAP AMONG DISEASES

Disease	Item Number																										
	7	8	9	17	19	22	25	30	31	32	34	39	48	51	52	53	54	56	65	66	67	70	102	103	104	105	106
	Percent "Yes"																										
350	59	56	56	22	26	22	11	26	04	15	33	15	22	11	15	11	15	19	19	26	00	30	06	23	10	03	00
471	09	07	09	58	49	60	47	56	37	23	53	02	33	15	15	09	03	00	28	21	09	19	02	33	26	09	09
360	03	06	12	33	24	12	12	73	64	70	61	88	30	24	36	12	27	06	18	24	09	24	09	18	15	09	03
527	09	05	06	18	15	10	11	40	13	15	25	12	65	63	61	57	74	64	06	10	03	25	09	24	20	07	06
724	17	04	08	25	25	14	16	29	14	12	26	14	21	14	21	11	13	14	36	54	30	47	07	43	43	08	07
630	17	09	13	18	13	11	10	23	21	14	24	14	24	19	24	14	22	14	13	22	08	22	49	69	61	31	25
Total Sample	13	09	09	18	17	13	10	26	17	14	25	13	27	19	24	13	19	10	13	20	06	19	11	29	29	08	06
	Significance Index																										
350	6	7	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
471	0	0	0	4	3	6	5	2	2	0	2	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
360	0	0	0	1	0	0	0	4	5	7	3	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
527	0	0	0	0	0	0	0	0	0	0	0	0	3	4	3	5	5	8	0	0	0	0	0	0	0	0	0
724	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	3	3	2	0	0	1	0	0
630	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	3	2	3	2

TABLE 7

SIGNIFICANCE INDICES FOR ITEMS WITH OVERLAP AMONG DISEASES IN MEN

Item Number	Item Content	Disease															Diseases Per Item
		020	121	184	360	370	382	388	440	460	471	495	496	500	501	504	995	
		Significance Index																
11	Lump in throat									3								1
12	Sneezing													5	3			2
13	Stuffed nose			5								3	3					3
14	Running nose													4	3			2
16	Frequent colds								3									1
17	Chest colds	3	4								4				4			4
19	Severe colds										3							1
20	Hay fever history													8	4			2
21	Asthma history													3	8			2
22	Constant coughing	5	4								6				5	3		5
23	Coughing blood	3																1
25	Chronic chest condition	4	3								5				4		3	5
26	Tuberculosis history	5															6	2
28	High blood pressure history					5												1

TABLE 7 (*Continued*)

30	Chest pains				4		4											2
31	Thumping heart				5		3	3										3
32	Rapid heart				7			3										2
33	Breathing difficulty						3								6	3		3
34	Shortness of breath				3										3			2
35	Shortness of breath at rest														5	3		2
39	Heart trouble history				8		5	3										3
45	Poor appetite		4															1
56	Stomach ulcer history		3															1
111	Fatiguability		3															1
125	Rheumatic fever history				5													1
132	Chronic disease history		3												3			2
	Items Per Disease	5	7	1	6	1	4	3	1	1	4	1	1	4	11	3	2	

TABLE 8

SIGNIFICANCE INDICES FOR ITEMS WITH OVERLAP AMONG DISEASES IN WOMEN

Item Number	Item Content	Disease										Diseases Per Item
		360	370	382	388	440	471	495	500	501	995	
		Significance Index										
12	Sneezing								7	5		2
13	Stuffed nose					3	3	3	5			4
14	Running nose						3	4	5			3
16	Frequent colds						4			3		2
17	Chest colds						8			5		2
18	In bed with colds	3				3	3					3
19	Severe colds					3	6			3		3
20	Hay fever history								7	5		2
21	Asthma history								3	8		2
22	Constant coughing						8			7		2
25	Chronic chest condition						8			3		2
26	Tuberculosis history										5	1
28	High blood pressure history		5	4								2

TABLE 8 (*Continued*)

30	Chest pains	4										1
31	Thumping heart	4										1
32	Rapid heart	5										1
33	Breathing difficulty	3								4		2
34	Shortness of breath	3										1
35	Shortness of breath at rest	4										1
39	Heart trouble history	8		3	3							3
41	Missing teeth			3								1
84	Faint feeling				3							1
92	Nail biting				3							1
118	Sickly person										3	1
125	Rheumatic fever history	8										1
132	Chronic disease history	3										1
133	Underweight										3	1
137	Injury history				3							1
181	Emotional instability				3							1
	Items Per Disease	10	1	3	5	3	8	2	5	9	3	

TABLE 9

SCORES FOR 6 PATIENTS.

Sex—	M M M F F F
Age—	66 35 26 22 32 32
Number of Complaints—	68 11 85 52 47 27

Disease	Score	Disease	Score
1	68 00 00 00 00 00	31	25 00 00 00 00 37
2	12 00 12 00 00 00	32	12 00 12 00 12 12
3	00 00 00 00 00 00	33	75 00 00 00 00 00
4	50 00 00 00 00 00	34	95 00 00 00 18 00
5	87 00 00 00 00 00	35	00 00 00 00 00 00
6	43 00 00 00 00 00	36	00 00 18 00 37 00
7	00 00 00 00 00 00	37	75 00 00 00 31 00
8	00 00 00 43 12 12	38	12 00 18 00 18 00
9	18 00 00 00 00 00	39	00 00 50 18 12 12
10	18 00 25 00 00 00	40	12 00 25 00 00 00
11	18 00 95 00 00 00	41	31 00 00 00 00 00
12	00 00 50 37 25 12	42	18 00 12 00 00 00
13	00 00 50 00 00 00	43	68 00 00 00 00 00
14	00 00 00 00 00 00	44	50 00 00 00 00 00
15	00 00 95 00 00 00	45	75 00 00 00 12 00
16	00 00 62 00 00 00	46	43 00 12 81 62 00
17	00 00 00 00 00 00	47	31 00 00 00 00 00
18	00 00 00 56 68 43	48	00 00 00 00 00 00
19	37 68 75 00 00 00	49	00 00 00 00 00 00
20	87 00 00 00 00 00	50	12 00 00 00 00 00
21	93 00 00 43 25 12	51	31 00 00 00 00 00
22	75 62 62 00 00 00	52	56 00 00 00 00 00
23	75 00 00 00 00 00	53	81 00 00 00 00 00
24	25 00 00 00 62 00	54	25 00 37 00 00 00
25	00 62 00 00 00 00	55	56 00 00 00 00 00
26	00 00 00 00 56 25	56	00 00 00 00 00 00
27	00 00 37 00 00 00	57	00 00 00 75 87 00
28	12 00 43 00 50 00	58	00 00 31 00 00 00
29	87 00 12 00 31 95	59	18 00 00 00 00 00
30	37 00 50 00 00 00	60	00 00 00 18 31 00

3. In Table 8, item 28 of disease 382 represents the co-existence of disease 370.

In the study of Brodman *et al.* [1959], the $\sum (S_{ik})_m^2$ model was used to diagnose diseases in the 1956 sample of patients. A constant critical score of 35 for all categories was adopted, after examination of the

scores obtained for the 1948–1949 sample. The use of critical scores determined for each disease independently may improve the model.

Table 9 gives the scores for 6 randomly selected patients diagnosed in the hospital as having rheumatic heart disease. The heading of the table gives the sex and age of the patients and the number of "yes" responses to the questionnaire. The number of diseases indicated for the 6 cases in the table ranges from 21 for the first case to 3 for the second and for the last case.

The model developed with the S_{jk} values derived from the 1948–1949 sample of patients was used to program a data processing machine (van Woerkom [1960]) and was tested with the 1956 sample of patients. The machine made correct assignments in 44 percent of cases where the 60 diseases were diagnosed in the hospital. In order to test for false positives, that is, diseases diagnosed by the machine when the disease was not present, a sample of 350 patients was selected for which the hospital records could give evidence for the presence or absence of a disease. In only 5 percent of cases did the model generate false positives. Results with the 350 patients are essentially the same as those obtained by a physician experienced in the interpretation of the CMI, except that for the diagnosis of psychoneurosis the physician made correct assignments more often than did the machine (81 percent as compared to 42 percent).

CONCLUSIONS

An approach used in efforts such as the one presented in this paper is best guided by the characteristics of the data interpreted. It is primarily because of the many false negatives in the large amount of data analyzed that the present model was developed. When small amounts of accurately defined data are analyzed the method may require revision. This paper presents a justification of the approach reported here, on an analytic basis and on the basis of the results obtained showing few false positives among many correct assignments.

BIBLIOGRAPHY

Brodman, K., Erdmann, A. J., Jr., Lorge, I., and Wolff, H. G. [1951]. The Cornell Medical Index—Health Questionnaire II. As a diagnostic instrument. *J.A.M.A 145*, 152–7.

Brodman, K., van Woerkom, A. J., Erdmann, A. J., Jr., and Goldstein, L. S. [1959]. Interpretation of symptoms with a data processing machine. *A.M.A. Archives of Internal Medicine. 103*, 776–82.

Brodman, K. [1960]. Diagnostic decisions by machine. I.R.E. *Trans. Medical Electronics. ME-7*, 216–9.

Crumb, C. B., Jr., and Rupe, C. E. [1959]. The automatic digital computer as an

aid in medical diagnoses. *Proceedings of the Eastern Joint Computer Conference*, 174–9.

Feller, W. [1957]. *Probability Theory and Its Applications*, 2nd. edition. New York, John Wiley and sons.

Ledley, R. S., and Lusted, L. B. [1959]. Reasoning foundations of medical diagnoses. *Science 130*, 9–21.

van Woerkom, A. J. [1960]. Program for a diagnostic model. *I.R.E. Trans. Medical Electronics. ME-7*, 220.

QUERIES AND NOTES

D. J. FINNEY, *Editor*

159 NOTE: Three-Quarter Replicates of 2^4 and 2^5 Designs

PETER W. M. JOHN

California Research Corporation, Richmond, California and University of California, Berkeley, California. U. S. A.

1. THE 2^4 DESIGN

The usual half replicate of the 2^4 design is defined by $1 = \pm x_1x_2x_3x_4$; the main effects are clear of two-factor interactions, but the two-factor interactions themselves are confounded in pairs. The three-quarter replicates given in this section enable both the main effects and two-factor interactions to be estimated clear of two-factor interactions.

The designs are obtained by omitting any one of the following quarter replicates, defined by $x_4 = \pm x_2$, $x_3 = \pm x_1x_2$.

I. 1, *ac*, *bcd*, *abd*; $x_4 = \quad x_2$, $x_3 = -x_1x_2$.
II. *a*, *c*, *abcd*, *bd*; $x_4 = \quad x_2$, $x_3 = \quad x_1x_2$.
III. *b*, *abc*, *cd*, *ad*; $x_4 = -x_2$, $x_3 = \quad x_1x_2$.
IV. *ab*, *bc*, *acd*, *d*; $x_4 = -x_2$, $x_3 = -x_1x_2$.

We consider the case in which I is omitted.

The three quarters may be combined into three half replicates; II, III, $x_3 = x_1x_2$; II, IV, $x_4 = x_1x_3$; III, IV, $x_4 = -x_2$.

We estimate A from the half replicate III, IV, where $x_1 = -x_1x_2x_4$, so that A is confounded with ABD. In the other half replicates A is confounded with the two-factor interactions BC or CD. The other effects are obtained, each from a half replicate, in the same way and the analysis is presented in tabular form below. It may be shown that these are the least squares estimates, when higher-order interactions are assumed to be negligible. BD, and the mean, may be estimated from two half replicates; the least squares estimates are then the averages of these two estimates.

2. THE 2^5 DESIGN

The main effects and two-factor interactions in the 2^5 experiment can all be estimated from a half replicate. The variance of each esti-

Effect	Sets Used	Confounding
A	III, IV; $x_4 = -x_2$	ABD
B	II, IV; $x_4 = x_1x_3$	$ABCD$
C	III, IV; $x_4 = -x_2$	BCD
D	II, III; $x_3 = x_1x_2$	$ABCD$
AB	II, IV; $x_4 = x_1x_3$	BCD
AC	III, IV; $x_4 = -x_2$	$ABCD$
AD	II, III; $x_3 = x_1x_2$	BCD
BC	II, IV; $x_4 = x_1x_3$	ABD
BD	II, IV; $x_4 = x_1x_3$	ABC
	or II, III; $x_3 = x_1x_2$	ACD
CD	II, III; $x_3 = x_1x_2$	ABD

mate is $\sigma^2/4$, and all the degrees of freedom are used for effects. We present in this section a three-quarter replicate in which all the effects are estimable with variance $3\sigma^2/16$ and which provides 8 degrees of freedom for error.

The three-quarter replicate using three quarters defined by $x_5 = \pm x_2x_3$, $x_4 = \pm x_1x_2x_3$ is discussed in Davies [1956]. In this design all the main effects and two-factor interactions, except E, AD, BC, are estimated by averaging the estimates from two half replicates with variance $3\sigma^2/16$. The latter three effects, however, are estimated from only a single half replicate with variance $\sigma^2/4$.

If the quarter omitted is defined by $x_4 = \pm x_1$, $x_5 = \pm x_2x_3$, we have a design in which each main effect and each two-factor interaction is estimated from two half replicates. The particular design obtained by omitting the quarter $x_4 = -x_1$, $x_5 = +x_2x_3$ has the additional advantage that it may be developed by augmenting a "one-at-a-time" experiment, Daniel [1958].

3. SOME ADDITIONAL DESIGNS

The above designs possess a certain symmetry inasmuch as each factor appears at its high level and at its low level the same number of times. Another set of designs may be obtained when a half replicate, in which the highest-order interaction is confounded with the mean, is augmented by a quarter which also confounds a main effect. For example, one might take three of the sets of treatment combinations defined by

$$1 = \pm x_1 = \pm x_2x_3x_4(x_5) = \pm x_1x_2x_3x_4(x_5).$$

In this design the first factor appears at one level in 2^{n-1} points and at the other level in 2^{n-2} points.

It may nevertheless be shown that the method of analysis given in the previous sections still gives the least squares estimates. Such a design has the advantage in the 2^4 case over the design given in Section 1 that the effect which is estimated with greatest precision is a main effect rather than a two-factor interaction.

REFERENCES

[1] Daniel, C., [1958]. On varying one factor at a time. *Biometrics 14*, 430–1.
[2] Davies, O. L., [1956]. *The Design and Analysis of Industrial Experiments*. Oliver and Boyd, London, 472–5.

160 NOTE: On the Extension of Stevens' Tables for Asymptotic Regression

S. Lipton
University of New South Wales
Kensington, Sydney, Australia.

Concerning H. Linhart's note ([1960]. *Biometrics 16*, 125), I think it is worthwhile pointing out my own extension to Stevens' tables. This work has been reported by H. D. Patterson ([1956]. *Biometrics 12*, 389 and [1958]. *Biometrika 45*, 323) and was carried out at Rothamsted Experimental Station. The extended tables give, in Stevens' notation, the six F's for $r = .10(.01).90$ for each integral value of n from 3 to 12 inclusive. These values do not in the main overlap Linhart's work.

The tables proved surprisingly "popular" and I received several requests for copies. However, as the entries corresponding to large r for small n seem to be somewhat inaccurate, (apparently they are correct only to about 4 significant figures compared to more than 6 figures elsewhere) I am now recomputing the values. I intend to take the opportunity of further extending the range of n up to 15.

161 NOTE: Corrected Error Rates for Duncan's New Multiple Range Test

H. Leon Harter
Aeronautical Research Laboratory
Wright-Patterson Air Force Base, Ohio, U. S. A.

The error rates published by the author [3] for Duncan's new multiple range test were based upon critical values tabulated by Beyer [1] and by Duncan [2], some of which were in error, as reported earlier by the author [4]. On the basis of the corrected critical values for this

CORRECTED TABLE 1C

MINIMUM VALUES OF α_L CORRESPONDING TO $\alpha_S = 0.05, 0.01$

$(\alpha_L)_{\min} \mid \alpha_S = 0.05$

m	$N = 2$	$N = 3$	$N = 4$	$N = 6$	$N = 10$	$N = 16$	$N = 25$	$N \to \infty$
2	.0500	.0500	.0500	.0500	.0500	.0500	.0500	.0500
3	.0496	.0443	.0425	.0411	.0402	.0396	.0395	.0391
4	.0463	.0395	.0372	.0354	.0342	.0338	.0334	.0329
5	.0429	.0357	.0334	.0315	.0301	.0298	.0295	.0290
6	.0400	.0329	.0305	.0287	.0275	.0269	.0266	.0261
7	.0375	.0305	.0283	.0265	.0254	.0248	.0245	.0239
8	.0354	.0286	.0265	.0247	.0236	.0231	.0228	.0223
9	.0335	.0270	.0250	.0233	.0222	.0216	.0214	.0209
10	.0320	.0256	.0237	.0221	.0210	.0206	.0203	.0199
12	.0294	.0236	.0216	.0202	.0192	.0188	.0185	.0181
14	.0274	.0219	.0202	.0188	.0179	.0175	.0172	.0168
16	.0257	.0206	.0189	.0176	.0168	.0164	.0162	.0158
18	.0244	.0196	.0180	.0167	.0160	.0155	.0153	.0149
20	.0233	.0186	.0172	.0160	.0153	.0148	.0146	.0142

$(\alpha_L)_{\min} \mid \alpha_S = 0.01$

m	$N = 2$	$N = 3$	$N = 4$	$N = 6$	$N = 10$	$N = 16$	$N = 25$	$N \to \infty$
2	.0100	.0100	.0100	.0100	.0100	.0100	.0100	.0100
3	.0098	.0085	.0081	.0077	.0075	.0074	.0074	.0073
4	.0089	.0072	.0067	.0063	.0061	.0060	.0059	.0058
5	.0079	.0063	.0058	.0054	.0052	.0050	.0050	.0049
6	.0071	.0056	.0051	.0048	.0045	.0044	.0044	.0043
7	.0065	.0050	.0046	.0043	.0041	.0039	.0039	.0038
8	.0059	.0045	.0041	.0039	.0037	.0036	.0035	.0035
9	.0054	.0042	.0038	.0036	.0034	.0033	.0033	.0032
10	.0050	.0039	.0035	.0033	.0031	.0030	.0030	.0029
12	.0044	.0034	.0031	.0029	.0028	.0027	.0027	.0026
14	.0039	.0031	.0028	.0026	.0025	.0024	.0024	.0023
16	.0036	.0028	.0025	.0024	.0023	.0022	.0022	.0021
18	.0033	.0026	.0024	.0022	.0021	.0020	.0020	.0020
20	.0030	.0024	.0022	.0020	.0019	.0019	.0019	.0018

CORRECTIONS TO TABLE 2

Comparison of Type II-III Error Rates for Single Classification

A. Values of $\beta_2' = (\beta_S')_{max}$ for $\alpha = 0.05$

N	m	$\delta'=1.5$	$\delta'=2.0$	$\delta'=2.5$	$\delta'=3.0$
5	3	.464	.198	.0602	
	4	.467	.197	.0577*	.0129
	5	.472	.200	.0572	.0120
	6	.478	.203	.0574	
	12	.508		.0626	
	20	.532		.0685	.0123
7	4	.286	.0723	.0111	

N	m	$\delta'=1.0$	$\delta'=1.5$	$\delta'=2.0$	$\delta'=2.5$
9	4	.535	.169	.0251	
	8	.577			

N	m	$\delta'=0.5$	$\delta'=1.0$	$\delta'=1.5$	$\delta'=2.0$
13	3			.0496	
	4			.0537	
	5		.372	.0572	
	6		.384	.0604	
	8		.403	.0659	
16	4	.775	.257	.0214	.00046
25	3		.0774		
	4		.0858		
	5		.0927		
	6		.0986		
31	4	.573	.0385		
	5		.0422		

*For this case $\beta_2' = (\beta_X')_{min} = .0818$

B. Values of $\beta_2' = (\beta_S')_{max}$ for $\alpha = 0.01$

N	m	$\delta'=2.0$	$\delta'=2.5$	$\delta'=3.0$	$\delta'=3.5$
5	3	.508	.229	.0726	.0183
	4	.487	.211	.0630	.0143
	5	.479	.204	.0588	.0124
	6	.476	.201	.0568	.0114
	12		.208	.0567	.0101
	16	.498	.215	.0589	.0103

N	m	$\delta'=1.5$	$\delta'=2.0$	$\delta'=2.5$	$\delta'=3.0$
7	3	.577	.234	.0556	.0088
9	3	.397	.0989	.0126	
	5	.411	.103	.0119	.00075

N	m	$\delta'=1.0$	$\delta'=1.5$	$\delta'=2.0$	$\delta'=2.5$
11	8		.302	.0470	
	12	.762	.324	.0528	
	16	.776	.341	.0577	.0033
13	8	.667	.201	.0185	.00051

N	m	$\delta'=0.5$	$\delta'=1.0$	$\delta'=1.5$	$\delta'=2.0$
16	5			.0897	.0035
	6	.932	.536	.0946	
	12	.947	.585	.116	
21	16	.929	.440	.0388	
41	3	.678	.0372		
	5		.0466		
	8	.750	.0565		
	12		.0668		
	16		.0743		
61	10	.588	.0057		

test recently published by the author [5], Table 1C of [3], which gives minimum values of the Type I error rate α_L for the *LSD* test corresponding to $\alpha_S = 0.05, 0.01$ for Duncan's test, has been completely recomputed, and the corrected table is given below. Also given is a list of corrections to Table 2 of [3], which gives values of $\beta_2' = (\beta_{\delta'})_{\max}$, the maximum combined Type II–III error rate for Duncan's test, for $\alpha = 0.05, 0.01$. To reduce the length of the list, corrections of a single unit in the last place have been omitted, except when there are larger corrections for other value (s) of δ' for the same values of N and m. [In both tables, N represents the size of each sample, and m represents the number of sample means being compared. In Table 2, δ' represents the ratio of the difference between the largest and smallest sample means in the group being tested to the pooled estimate s of the population standard deviation σ]. Table 3 of [3], which gives maximum required sample sizes $(N_S)_{\max}$ for Duncan's test for prescribed values of α, β, and m, has not been recomputed on the basis of the corrected critical values. If this were done, there might be an occasional change of 1, but most values are correct as they stand.

REFERENCES

[1] Beyer, William H. [1953]. Certain percentage points of the distribution of the studentized range of large samples. Virginia Polytechnic Institute Technical Report No. 4.

[2] Duncan, David B. [1955]. Multiple range and multiple F tests. *Biometrics 11*, 1–42.

[3] Harter, H. Leon [1957]. Error rates and sample sizes for range tests in multiple comparisons. *Biometrics. 13*, 511–36.

[4] Harter, H. Leon. [1957]. Critical values for Duncan's new multiple range test (abstract), *Jour. Amer. Stat. Assoc. 52*, 372.

[5] Harter, H. Leon. [1960]. Critical values for Duncan's new multiple range test. *Biometrics 16*, 671–85.

162 QUERY: Multiple Comparisons between Treatments and a Control

For multiple comparisons between p treatments and one control Dunnett [1955] has tabulated values of t from which confidence limits can be calculated. His tables provide for values of p from 1 to 9, and for degrees of freedom from 5 to infinity. The full tables allow for different numbers of tests (n) for each treatment, but when the number is the same for each treatment the number of degrees of freedom is given by $(p + 1)(n - 1)$, so that a simpler table can be derived for p against n, including the following:–

n	$P = 0.05$				$P = 0.01$			
	$p = 2$	4	7	9	$p = 2$	4	7	9
2		2.85	2.81	2.81		4.43	3.96	3.83
3	2.34	2.47	2.56	2.60	3.61	3.45	3.39	3.38
4	2.18	2.36	2.48	2.54	3.19	3.20	3.22	3.24

Does it really make sense for an increase in the number of treatments in each test to lead to smaller significant differences when there are only two tests of each, but larger when there are four or more? Or, when there are three tests of each, for an increase in treatments to increase the size of differences which are significant at $P = 0.05$ but reduce those at $P = 0.01$?

REFERENCE

Dunnett, C. W., [1955]. A multiple comparison procedure for comparing several treatments with a control. *J. Amer. Stat. Assoc. 50*, 1096.

ANSWER:

At first sight the behaviour of the table does appear rather odd. Intuitively one would probably expect an increase in the number of treatments to lead to a larger significant difference in all circumstances. This would certainly be the case if the only error variation were that associated with the control and treatment means. But it must be remembered that the variance has to be estimated, and the value of t has to allow for error in its value too. The reason the value of t decreases with p for the small values of n is that there the number of degrees of freedom for estimating the variance is small and, in effect, we gain more from the increase in degrees of freedom than we lose as a result of the greater number of treatment means.

If the table were extended to larger values of p, the value of t for $n = 2$, and for $n = 3$ when $P = .01$, would pass through a minimum and then increase with the value of p, indicating that a point is reached at which any further increase in the number of degrees of freedom is not sufficient to compensate for the increase in the number of treatments. For values of $n \geq 4(P = .01)$ and $n \geq 3(P = .05)$, this is the situation which pertains for *all* values of p.

It may be verified that the same phenomenon occurs in the case of the Studentized range, if the percentage points of that distribution are tabulated with n as parameter instead of the number of degrees of freedom.

C. W. DUNNETT
Lederle Laboratories Division
American Cyanamid Company
Pearl River, New York

163 QUERY: Error Rates in Multiple Comparisons

Many biologists, realising the necessity for statistical examination of data, are unfortunately without access to professional advice, and must perforce search the literature themselves. One such worker encountered Steel's description [1959] of a multiple comparison sign test, which seemed to fill an urgent need, yet on closer study to reveal an apparent paradox. As his example Steel applies his test to data from 13 soybean yield trials of 3 varieties, and concludes that one variety (Z) was significantly better than control (C) because it was inferior ($-$) to control in only 2 of the trials. If, however, there had been five varieties in each trial, Z would not have been significantly different from control, because with 5 treatments the number of minus signs must not exceed one for significance. And yet the inclusion of two more varieties in the trials could not have made one iota of difference to the yield figures for Z and C from which alone the number of inferiorities was obtained. How can the presence or absence of data about W, Y, etc., entirely irrelevant to the $Z : C$ comparison and to all the computations involved alter the probability that Z is the same as C? Or are there implied assumptions, not apparent to the amateur reading Steel's paper, about the behaviour of W and Y relative to Z and C which determine when this test should be used in preference to Dixon and Mood's [1946] simple sign test?

REFERENCES

Dixon, W. J., and Mood, A. M., [1946]. The statistical sign test. *J. Amer. Stat. Assoc. 41*, 557.

Steel, R. G. D., [1959]. A multiple comparison sign test: treatment versus control. *J. Amer. Stat Assoc. 54*, 767.

ANSWER:

This query seems to arise from a failure to recognize that several definitions of error rate exist. These, of which only two are considered here, were introduced in connection with multiple comparison procedures and are just now finding their way into application and textbook. The decision as to which error rate is more appropriate and what significance level is to be used must, finally, be made by the experimenter.

Error rates are defined on the basis that the null hypothesis is true.

$$\text{Per-comparison error rate} = \frac{\text{Number of comparisons falsely declared significant}}{\text{Total number of comparisons}}.$$

Here, the comparison is the conceptual unit. When comparisons are chosen without reference to the experimental results, a test such as the (parametric) *lsd* provides a per-comparison error rate at the tabulated value regardless of whether or not the comparisons are independent. A criticism of tests using this error rate is that, if many comparisons are made, the probability of finding at least one falsely significant difference is high.

$$\text{Experimentwise error rate} = \frac{\text{Number of experiments with at least one difference falsely declared significant}}{\text{Total number of experiments}}.$$

Here, the experiment is the conceptual unit. For pairwise comparisons, we agree to test the largest difference when it occurs. If this difference is significant, then this experiment is one for which at least one difference has been declared significant. Tukey's (parametric) test based on the studentized range has such an error rate; Dunnett's (parametric) test of treatments against control also has this error rate.

Clearly, tests with an experimentwise error rate will require a larger difference for significance than that required by the *lsd*. Furthermore, the required difference will increase with the number of treatments. This becomes a basis of criticism: the probability of detecting a real difference of fixed size is lower when an experimentwise error rate is used and decreases with the number of treatments.

Error rates are not, perhaps, as different as they first seem. If α_C and α_E represent per-comparison and experimentwise error rates respectively, then for k independent comparisons $(1 - \alpha_E) = (1 - \alpha_C)^k$. For dependent comparisons, the relation is not as obvious. Harter

[1957] has tabulated comparisons of significance levels for a number of multiple comparisons test procedures intended to compare all pairs of means.

It is apparent, then, that if an investigator chooses a definition of error rate, a test, and a significance level, then this information permits him to compute the significance level for any other suitable test and error rate. This knowledge eliminates any paradox.

No statistical assumptions determine which test is appropriate, regardless of whether the test is parametric or non-parametric. One chooses an error rate and significance level, prior to the conduct of the experiment, to give the kind of protection desired. It seems to me that the basic problem concerns a power function, presumably related to "urgent need" and, in turn, the sample size necessary to accomplish clearly stated aims.

REFERENCE

Harter, H. Leon, [1957]. Error rates and sample sizes for range tests in multiple comparisons. *Biometrics 13,* 511–36.

R. G. D. Steel
Institute of Statistics
North Carolina State College
Raleigh, N. C., U. S. A.

BOOK REVIEWS

J. G. SKELLAM, *Editor*

Members are Invited to Suggest Books for Listing or Review to the Editor

10 FALCONER, D. S. **Introduction to Quantitative Genetics.** Edinburgh and London: Oliver & Boyd Ltd., 1960. Pp. ix + 365. Figures and Tables. 35s.

W. F. BODMER, *Department of Genetics, University of Cambridge, Cambridge, England.*

There has long been a need for a good introductory textbook on quantitative genetics, and Dr. Falconer's new book satisfies this need more than adequately. The term *quantitative genetics* is here used in its broadest sense to include both the study of genetic changes in populations and the study of the inheritance of continuous variation. The genetic analysis of changes in quantitative characters in both natural and artificial populations requires a close interplay between these two aspects.

The first five chapters of the book are concerned with the analysis of changes in gene frequency in natural and artificial populations. These provide perhaps the best elementary exposition I know of what is usually called *population genetics*. Dr. Falconer has a great gift for lucid and stimulating exposition. He presents the essential mathematics in a way that should be intelligible event to the least mathematically minded biologist. The numerical examples are excellently chosen to help in the understanding of the mathematical formulae as are also the graphical illustrations. In my opinion Dr. Falconer rightly assumes a knowledge of elementary genetics and later also of the elements of statistics. His book would have been unnecessarily burdened by any attempt to make it complete in these respects.

The genetics of continuous variation is treated in the following fifteen chapters with the same high standard of exposition. This subject has been approached in a variety of ways by different research schools and suffers from a confusion of terminology and notation. Dr. Falconer sets himself the difficult task of integrating the subject at an elementary level and is on the whole most successful. It is not surprising if the subject matter is somewhat biased toward the approach of the Edinburgh 'school', of which he is a member. Dr. Falconer is at least honest in delineating controversial topics and making it quite clear when a discussion is more a matter of personal opinion and preference than of established fact.

Any criticism of the subject matter is perhaps more a criticism of the development of the subject than of Dr. Falconer's account of it. However there is a notable lack of any emphasis on the importance of linkage, and the treatment of Mather's concept of *balance* is rather brief. The importance of epistacy is also, in my opinion, underestimated. A good epistatic genotype may be quite rare, and so contribute only little to an interaction variance, and yet be of considerable importance to the practical breeder. As Dr. Falconer says in his final discussion, there is need for much more experimental work. This should not, I feel, be directed at estimating more and more heritabilities and variance components for a wide selection of characters and organisms, as he suggests, but aimed at a more penetrating and specific analysis of the fundamental mechanisms underlying the inheritance of quantitative characters.

In conclusion, this is a book that should appeal to a wide audience. For the experienced geneticist it provides a stimulating and refreshing review of the subject, and for the newcomer to genetics it offers an excellent introduction. There is no doubt that it will be the best standard textbook of its kind for some years to come.

11 GOLDBERG, S. **Probability: An Introduction.** Englewood Cliffs, N. J.: Prentice Hall, Inc. 1960. Pp. xiv + 322. 30 Figures, 44 Tables. $5.95.

F. N. David, *Department of Statistics, University College, London, England.*

This book, written by a pupil of W. Feller, has been tried out, we are informed, on the freshman mathematics class of Oberlin College. This should give some idea of the level of mathematical sophistication which is aimed at and which the author generally achieves.

The first chapter on the manipulations and the algebra of sets is necessary for the main development of the book. It will be found difficult for persons wishing to learn probability with a view to practical applications, but the would-be mathematician will find it interesting, for the elements of mathematical logic are introduced. Probability in finite sample spaces follows. The author does not really attempt to justify the calculation of a probability, being concerned simply to define a one to one correspondence between the field of events and the probability set, and then to discuss the sub-sets involved. He says: "The theory of probability begins when a sample space, S, the mathematical counterpart of an experiment, is specified"—an entirely proper point of view for a mathematican to take. The second chapter, which covers conditional and compound probabilities, Bayes' formula and independence, will be found rather stiff reading to the uninitiated.

After a short discursus into elementary enumeration through the binomial and multinomial coefficients, with the arithmetic triangle and several combinatorial theories expounded, the remainder of the book is concerned with the properties of random variables. These are few brief notes towards the end on testing a statistical hypothesis and on decision-making.

For those persons who already know some probability theory this book will be enjoyable reading, if only to appreciate how familiar theorems can be wrapped up in mathematical language. For mathematicians in embryo who want to learn how a mathematician looks at the elements of probability theory this book will serve a useful educational purpose. For the statistician in embryo this book is not suitable since enlightenment can be achieved more simply elsewhere. (The reviewer has in mind, for example, *An Introduction to Probability Theory and its Applications* by W. Feller). But possibly the author did not intend to cater for this last class anyway.

ABSTRACTS

The following are abstracts of papers presented at meetings of the British Region held on October 27 and December 13, 1960.

745 W. T. WILLIAMS (Botany Department, University of Southampton). **Some Applications of Electronic Computers in Plant Ecology and Taxonomy.**

In ecological survey the population under study consists of sample areas (quadrats) defined by the presence or absence of a number of plant species, whose names and ecological requirements may not be known. The problem is to find the major discontinuities in the population. This can, in theory, be solved by successively subdividing the population so as to reduce the associations between the species by the greatest possible extent at each stage; this in turn involves setting up the correlation matrix at each stage and finding the species with the highest loading on the first principal axis of this matrix. To hold the entire matrix in a computer reduces unacceptably the space available for data. Instead of r, therefore, $|r|$ is calculated, and the $|r|$ values for each species summed as they are obtained. The species with the highest sum of $|r|$ is necessarily that with the highest loading on the first centroid axis of the matrix ($|r|$); this has been found to be an acceptable approximation to the answer required. Examples will be given of the application of this method to actual areas.

Valuable information on the structure of plant communities can be obtained by carrying out the corresponding inverse (Q-technique) subdivision; and, by collating the direct and inverse analyses, a completely objective account of the vegetation of a complex area can be given in extremely economical terms.

In taxonomy it is sometimes desirable to carry out a conventional factor anlysis, with communalities and rotation. No programmes of this type seem to be available in Britain; but the standard programmes for the extraction of latent roots and vectors greatly reduce computational labour. Moreover, the existence of general "matrix schemes" makes calculation of specification equations a relatively simple task.

746 J. N. R. JEFFERS (Forestry Commission, Wrecclesham, Farnham, Surrey). **Data Processing in Biological Research.**

The first task in the application of electronic digital computers to the statistical analysis of data arising in biological research has been the writing of the programmes for the actual analysis. This is, however, only a small part of the whole process, and the paper presents the results of recent work on the recording, compilation, and storage of data on paper tape and magnetic tape. This work suggests the need for a complete revision of the methods used in handling biological data, and has shown very large savings in the cost of research and in the unproductive use of key research workers. Also considered is the problem of the output results from the computer in a directly intelligible form, though progress has been slower in this field.

747 F. YATES (Rothamsted Experimental Station, Harpenden, Herts.). **Problems Arising in the Use of Electronic Computers in Statistical Analysis.**

A small electronic computer has been in use in the Rothamsted Statistics Department for the last six years. It has been used for a wide variety of statistical problems arising in agricultural and biological research. It was early applied to the analysis of replicated experiments and a large number of such analyses are now performed every year. More recently its use in the analysis of surveys has been developed. A general programme which has been written for survey analysis will be outlined and used to illustrate the types of problem that are encountered in using electronic computers for statistical work.

748 C. W. EMMENS (University of Sydney, N. S. W., Australia and Royal Veterinary College, London). **The Planning and Analysis of Some Field Trials with Cattle.**

Three field trials of the fertility of deep frozen semen have recently been conducted in Sydney. They were as far as possible completely balanced, on a factorial scheme, and have given remarkably consistent results, with residual χ^2-values within expected limits. This is attributed to the care with which randomisation and technique were controlled under difficult conditions.

749 C. C. SPICER (Imperial Cancer Res. Fund, Lincoln's Inn Fields, London). **Problems in the Analysis of a Large Scale Clinical Trial.**

The purpose of this trial was the evaluation of chemotherapy in the treatment and prevention of chronic bronchitis. It presents a number of points of interest both in planning and analysis.

The following are abstracts of papers presented at a special meeting on "BIOMETRICAL ASPECTS OF PLANT GROWTH" held in London on January 2, 1961, by the British Region in conjunction with the Society for Experimental Biology.

750 J. A. NELDER (National Vegetable Research Station, Wellesbourne, Warwick). **Models and Experiments for Growth Analysis.**

The relationship of growth analysis to other forms of statistical analysis, particularly those of long-term experiments and multivariate analysis, is indicated. The necessity for the development of methods for analysing the inter-related growths of different parts of a plant is stressed, and a first effort model of differential equations analagous to linear regression is proposed. Types of deviation from the assumed model, both systematic and random, and the kind of experiment needed to detect them are discussed together with problems arising in the estimation of parameters in growth equations. Possible applications of the model are indicated.

751 S. C. PEARCE and C. S. MOORE (East Malling Research Station, Nr. Maidstone, Kent). **A Study of the Sources of Variation in Growth of Fruit Trees.**

A method of studying growth previously described (*Biometrics 16*, 1–6) is applied to annual records of apple tree trunk circumference. The purpose of the investigation is to determine the main causes of variation in field trials with clonal fruit trees. Three standard errors are plotted against the logarithmic development of girth, namely, (a) log (girth), (b) increment in log (girth) since planting, and (c) log (girth) adjusted by girth at planting. A study of these curves enables the total variation resulting after a period of growth to be divided between initial variation, and initial growth rate negatively correlated with initial size differences, and a subsequent growth rate which is uncorrelated with initial size. Comparison of groups of trees having different treatments demonstrates the importance of good growing conditions in minimising variation. The reduction in positional variation by removing block effects and the results of cropping are also considered. The effects of these modifications in conditions accord with the theoretical model. Some measures are suggested for minimising error in field trials.

752 M. J. R. Healy (Rothamsted Experimental Station, Harpenden, Herts.). **Experiments for Comparing Growth Curves.**

Consider an experiment designed to study the differential effects of treatments on the growth curve of some organism. Suppose that the experimental subjects are allocated to the treatments by some process of randomisation and that observations are taken when the treatments are first applied and at regular intervals thereafter. Wishart suggested that tests of significance for the existence of treatment effects could be made by fitting a polynomial to the data from each subject and by applying standard statistical techniques to the coefficients of the polynomials. The paper discusses Wishart's technique and extends it to provide estimates of the treatment effects.

A general discussion was led by F. L. Milthorpe (University of Nottingham.)

CORRECTION

E. N. Hey and M. H. Hey [1960]. The Statistical Estimation of a Rectangular Hyperbola. *Biometrics 16*, 606–17.

The expression, nine lines from the bottom of page 609, should coincide with the column vector on the right of the matrix equation at the top of page 610. The solution for a, b, and $c - ab$ on page 610 should have coefficients bracketed to agree with the matrix equation at the top of the page. In footnote 2, page 615, Ey in the matrix should be E_y and on page 617, line 6, χ should be replaced by x.

THE BIOMETRIC SOCIETY

International

ANNUAL FINANCIAL STATEMENTS

THE BIOMETRIC SOCIETY 1960

BALANCE SHEET

Assets

Cash: Bank Balance	$9,605.90	
Petty Cash	19.40	$9,625.30
Liabilities		
Subscriptions, 1961	$ 107.50	
Dues, 1961	54.50	$ 162.00
Surplus, Jan. 1, 1960 (including petty cash)	7,048.64	
Gain for Period	2,414.66	9,463.30
		$9,625.30

Audited: B. G. Selvidge (Signed)
Date: March 20, 1961

INCOME AND EXPENDITURE STATEMENT

Income

Subscriptions, 1959	$ 358.75	
Subscriptions, 1960	6,225.00	$ 6,583.75
Dues, 1959	105,50	
Dues, 1960	2,466.00	2,571.50
Sustaining memberships, 1960		650.00
Back dues and subscriptions		32.00
Regional allotments	116.50	
BIOMETRICS allotment from sustaining members	250.00	
Back issues	93.50	
Member subscriptions to Journal of ASA	35.00	
Sale of directories	2.00	
Overpayments	20.50	
Printing—Harvard University Press	30.00	
Advance to Pharmacology Symposium	1,000.00	
Bank refund of service charge	1.00	
	1,548.50	

Less: Credits and regional allotments used	138.00	1,410.50
Total Income		$11,247.75

Expenditures

BIOMETRICS	$6,996.28	
Postage	250.23	
Office supplies	2.89	
Printing	260.29	
President's Office	200.00	
Member subscriptions to Journal of ASA	35.00	
Addressing and mailing services	88.40	
Pharmacology Symposium	1,000.00	$ 8,833.09
Excess of Income over Expenditures		2,414.66
		$11,247.75

Audited: B. G. Selvidge (Signed)
Date: March 20, 1961

SECRETARY'S ACCOUNT

STATEMENT OF INCOME AND EXPENDITURE

For the Period 1st January, 1959 to 31st December, 1960

Income

Balance in hand, 31st December, 1958	£ 64.10. 4d
Received from Treasurer	266.11. 3d
Menbership dues received	1.12. –d
	£332.13. 7d

Expenditures

Office equipment and stationary	£ 56. 2. 6d
Secretarial assistance	150. –. –d
Printing and binding	13.10. 7d
Postage	49.12. 1d
Travel	29.15. 9d
Books	18. 8. –d
Membership dues paid over	1.13. 6d
	£319. 2. 5d
Balance, 31st December, 1960	£ 13.11. 2d

I certify the above to be a true record of my transactions on behalf of the Biometric Society

Signed—.....
M. J. R. Healy
Secretary

I have examined the account book and vouchers produced by the Secretary and certify that the above statement is in accordance therewith.

Signed—.....
E. Church, A.A.C.C.A.

BIOMETRICS, VOLUME 16

STATEMENT OF OPERATIONS

For the Year Ending January 31, 1961

Income			
Biometric Society:			
907 Subscriptions at	$ 4.00	$ 3,628.00	
983 Subscriptions at	2.75	2,703.25	
9 Sustaining Members at	25.00	225.00	$ 6,556.25
1161 Direct Subscriptions at	7.00		8,127.00
Sale of Back Issues:			
Biometric Society		$ 374.00	
Editor's Office		4,474.00	4,848.00
March 1960 Issues at Cost to Biometric Society			61.28
Sale of Reprints			1,537.75
Over Payments			11.50
Payments to Regions			6.50
Total Income from Operations			$21,148.28
Expenditures			
Cost of Printing Journals			
Issue No. 1	$2,928.85		
Issue No. 2	3,344.71		
Issue No. 3	2,988.44		
Issue No. 4	4,467.72	$13,729.72	
Mailing and Express Charges			
Issue No. 1	$ 206.28		
Issue No. 2	229.82		
Issue No. 3	310.70		
Issue No. 4	346.56	1,093.36	$14,823.08
Cost of Printing Reprints (See Comments)			
Issue No. 4 (1959)	$ 290.69		
Issue No. 1	329.43		
Issue No. 2	331.79		
Issue No. 3	370.75	$ 1,322.66	
Mailing Charges Reprints (See Comments)			
Issue No. 4 (1959)	$ 24.01		
Issue No. 1	27.99		
Issue No. 2	22.75		
Issue No. 3	38.94	113.69	$ 1,436.35

Operating Expenses			
Stamps	$ 297.28		
Office Supplies	523.68		
Mailing Lists and Addressographing	131.75		
Back Issue Moving Costs	182.94		
Back Issue Handling and Storage	329.76		
Back Issue Insurance	46.35		
Auditing	45.00		
Bank Exchange and Service Charges	2.43		
Refunds and Overpayments on Subscriptions	55.25		
Salaries and F.I.C.A. Taxes	525.02		
Transfers to Regions	6.50		
Express Charges	13.97	$ 2,159.93	
Total Expenditures from Operations			18,419.36
Surplus from Operations			$ 2,728.92
Non-Operating Items			
Income			
Bank Interest	$ 680.75		
Refund on Insurance Policy	31.66		
Bank Credit and Exchange	2.06		
Announcement	57.00		
Refund on Blacksburg Postage Deposit	66.61		
Correction of Surplus from Previous Volumes (See Comments)	28.35	$ 866.43	
Expenditures			
Credits of 1959 Accounts Receivable		33.50	
Surplus from Non-Operating Items			832.93
Gross Surplus, 1-31-61			$ 3,561.85

BIOMETRICS, VOLUME 16
BALANCE SHEET
January 31, 1961

Assets		
Accounts Receivable	$ 2,089.25	
Bank Balances		
Savings Account #1 (Blacksburg)	10,397.02	
Savings Account #2 (Blacksburg)	9,637.50	
Savings Account #3 (Blacksburg)	3,168.41	
Checking Account (Tallahassee)	10,875.78	
Prepaid Insurance (Expires 4-25-63)	139.05	
Total Assets		$36,307.01

Liabilities and Surplus

Subscriptions to Volume #17	$ 5,993.75	
Subscriptions to Volume #18	182.00	
Subscriptions to Volume #19	42.00	
Subscriptions to Volume #20	7.00	
Surplus from Previous Volumes	26,520.41	
Surplus from Volume #16	3,561.85	
Total Liabilities and Surplus		$36,307.01

Comments

(1) Not included in Assets is stock of back issues from Volumes 1–16.
(2) Not included in Expenses is printing bill for December reprints of $507.20
(3) Corrections of surplus from previous volumes results from a reconciliation of subscription card files and actual liabilities due to prepaid subscriptions.

Audited February 17, 1961
Richard Q. Conrad (Signed)

Region Belge

L'Assemblée Générale annuelle de la Société Adolphe Quetelet s'est tenue à Bruxelles, dans la salle du Conseil de la Bibliothèque Royale de Belgique, le 25 février 1961. Le Conseil d'Administration pour 1961 et 1962 a été constitué comme suit:

Président: Professeur Maurice WELSCH.
Vice-Présidents: Prof. R. CONSAEL, H. LAUDELOUT, A. LECRENIER, J. REUSE et A. VAN DEN HENDE.
Secrétaire: Dr. L. MARTIN.
Secrétaire-adjointe: Anne LENGER.
Trésorier: Pierre GILBERT.
Membres: Dr. H. G. LION,
Mme OSLET-CONTER,
Mr. H. ROGGEN.

L'Assemblée Générale a été suivie d'une réunion organisée conjointement par la Société Quetelet, le Comité belge d'histoire des Sciences et la Société belge de Statistiquë Les exposés suivants ont été donnés:

"*Les rapports historiques entre l'anthropologie et la statistique*" par Monsieur A. LEGUEBE;
"*La distribution des caractères anthropométriques*" par le Professeur C. GINI, Président de l'Institut International de Sociologie à Rome.

British Region

At a meeting held on February 28th, 1961, the following papers were read and discussed:

P. D. Oldham—The Distribution of Arterial Pressure in the General Population.
L. R. Taylor—A Power Law Transformation for Aggregated Populations.

W.N.A.R.

TREASURER'S REPORT—DECEMBER 31, 1960

Bernice Brown, *Treasurer WNAR*

Receipts

Balance forwarded from previous treasurer, Dec., 1959	$ 461.22
Received from 1960 dues (143 @ 7, 10 @ 4)	
1961 dues (61 @ 7, 1 @ 4)	1472.00
Received for orders of back issues of *BIOMETRICS*	16.00
Received from Biometric Society (credit for one institutional membership)	10.00
Received from Biometric Society Treas. (credit for 4 transfers from ENAR)	4.00
Received from ENAR for information circulars	41.60
Total	$2004.82

Expenses

To Treasurer International (143 @ 6, 10 @ 4)	898.00
To *BIOMETRICS* for back issues ordered by members	16.00
To Marion Sandomire—mailing expense	11.15
To printers for stationary and supplies	111.84
To C. Zippin—partial expense to meeting	26.18
To W. Becker—mailing expense	12.00
To ENAR Treasurer—mailing expense	5.00
To printers for notices of meetings	45.40
To Pullman Herald for stationary	92.08
Total	$1217.65
On hand—checkbook balance	$ 787.17
Total	$2004.82

NEWS AND ANNOUNCEMENTS

Members are invited to transmit to their National or Regional Secretary (if members at large, to the General Secretary) news of appointments, distinctions, or retirements, and announcements of professional interest.

Robert W. Allard has been granted a Guggenheim Fellowship to study stochastic processes in genetics during a six months sabbatical leave to be spent at Oxford University in the Biometric Unit beginning February 1, 1961.

Neeti R. Bohidar has accepted the position of Assistant Professor of Statistics in the Department of Applied Statistics, Statistical Laboratory, Utah State University, Logan, Utah. He received his Ph.D. degree in Statistics from Iowa State University in 1960.

Joseph G. Bryan, formerly Senior Operations Research Analyst at the Central Research Laboratory, American Machine and Foundry Co., recently changed positions and is now employed by Travelers Weather Research Center, Inc., 11 Newport Avenue, West Hartford 7, Connecticut.

William R. Buckland has resigned from London Transport Executive to take charge of the new Statistical Advisory Service of The Economist Intelligence Unit Ltd., 22, Ryder Street, London, S.W.1.

Lyle D. Calvin, Oregon State University, spent the 1960 fall term on sabbatical leave with the Department of Statistics, Harvard University.

Douglas G. Chapman participated in the Fourth Annual Meeting of the North Pacific Fur Seal Commission held in Tokyo, January 20 to February 4 as a Scientific Advisor to the U. S. Commissioner. The North Pacific Fur Seal Commission is composed of representatives from Canada, Japan, U.S.S.R., and the U.S.A.

Franklin A. Graybill recently left Oklahoma State University to become Professor of Mathematical Statistics and Director of the Statistical Laboratory at Colorado State University, Fort Collins, Colorado.

Iowa State University granted Doctor of Philosophy degrees, with major in Statistics, to the following people during the Winter Quarter graduation: Roger McCullough, Jose Nieto de Pascual, J. N. K. Rao and Thomas Neil Throckmorton.

Gwilyn Jenkins, after spending a year at Stanford, is now back at his previous post as lecturer in Mathematical Statistics, University of London.

Kuo Hwa Lu, formerly Associate Professor, Department of Statistics, Utah State University, has taken the position of Associate Professor of Biostatistics at the University of Oregon Dental School, Portland.

Lincoln E. Moses, Professor of Statistics at Stanford University, is spending a year on sabbatical leave at the Department of Social Medicine, Oxford University, England. Dr. Moses received a Guggenheim Fellowship.

Donald B. Siniff had taken a position as Biometrician in the Division of Biological Research, Alaska Department of Fish and Game. He previously was at North Carolina State College working under the Southeastern Cooperative Wildlife Statistics Project.

TRAVEL GRANTS FOR ATTENDANCE AT THE INTERNATIONAL CONGRESS OF MATHEMATICIANS

Travel grants will be made to a number of mathematicians who wish to attend the International Congress of Mathematicians in Stockholm, on August 15–22, 1962. It is hoped that funds available through various sources may provide travel assistance for a considerable number of mathematicians.

There will be a greater effort than in the past to give aid to younger people. As grants will be made only to those who have filed applications, it is urgent that any who wish to receive a grant should fill out and file an application. Younger people are urged to file applications so that their cases can be considered. Applications can be obtained from the Division of Mathematics, National Academy of Sciences, National Research Council, Washington 25, D. C. by requesting an application for a travel grant to the 1962 International Congress.

The deadline for filing of applications is November 1, 1961, and an attempt will be made to announce the grants by January 1, 1962.

Awarding of grants will be made only to those persons whose applications have been received, in good order, by November 1. The selection will be made by a committee consisting of the regular Committee on Travel Grants of the Division of Mathematics of the National Academy of Sciences—National Research Council enlarged to include representatives of societies affiliated with the Division and representatives of various governmental agencies.

TRAINEESHIPS FOR PUBLIC HEALTH STATISTICIANS

The Public Health Service has announced the availability of traineeships for graduate training of professional public health personnel during the 1961–1962 academic year.

Traineeships in public health statistics are available to qualified persons. They provide stipends from a minimum of $250 per month for a post-bachelor candidate to a maximum of $400 per month for a post doctoral candidate and additional allowances for dependents, travel of the trainee, and academic tuition and fees.

Additional information and application forms may be secured from the Division of General Health Services, Public Health Service, U. S. Department of Health, Education, and Welfare, Washington 25, D. C.

GRADUATE TRAINEESHIPS IN BIOMETRY

Training programs designed to prepare students in the application of statistical and mathematical methods to biological problems, particularly those related to health and medical sciences, now exist in more than 20 universities throughout the country. Supported by training grant funds from the Public Health Service, NIH, these programs provide unusual opportunities for careers in teaching, research, and consultation. Employment opportunities for biometricians are excellent, with the demand by governmental and voluntary health agencies, medical research and educational institutions, and industry running far in excess of the available supply of trained personnel.

Programs of study are individually designed to lead to doctoral degrees, and in special instances, to other academic degrees. Traineeship stipends are provided at various levels depending on previous education and experience of the trainee and include allowances for dependents. Substantially full economic support or partial

support may be provided, depending upon the proportion of time spent in training.

Interested applicants are encouraged to correspond with one of more of the Program Directors listed below because course offerings, as well as specific research problems for application of learned skills, vary from school to school.

Dr. Virgil Anderson
Purdue University
Lafayette, Indiana

Dr. George F. Badger
Western Reserve University
Cleveland 6, Ohio

Dr. Jacob E. Bearman
University of Minnesota
Minneapolis 14, Minnesota

Dr. Antonio Ciocco
University of Pittsburgh
Pittsburgh 13, Pennsylvania

Dr. Wilfrid J. Dixon
Medical Center, U.C.L.A.
Los Angeles 24, California

Dr. W. T. Federer
Cornell University
Ithaca, New York

Dr. John W. Fertig
Columbia University
New York 32, New York

Dr. Franklin A. Graybill
Colorado State University
Fort Collins, Colorado

Dr. Bernard G. Greenberg
University of North Carolina
Chapel Hill, North Carolina

Drs. John Gurland and T. A. Bancroft
Iowa State University
Ames, Iowa

Dr. Boyd Harshbarger
Virginia Polytechnic Institute
Blacksburg, Virginia

Dr. Allyn Kimball
Johns Hopkins University
Baltimore 5, Maryland

Dr. Schuyler G. Kohl
State Univ. of N. Y., Col. of Med.
Brooklyn 3, New York

Dr. Robert F. Lewis
Tulane University
New Orleans 12, Louisiana

Dr. Eugene Lukacs
The Catholic University of America
Washington, D. C.

Dr. Paul Meier
University of Chicago
Chicago 37, Illinois

Prof. Felix Moore
University of Michigan
Ann Arbor, Michigan

Dr. Lincoln E. Moses
Stanford Medical School
Stanford, California

Dr. Hugo Muench
Harvard School of Public Health
Boston 15, Massachusetts

Dr. Robert Quinn
Vanderbilt University
Nashville 5, Tennessee

Prof. J. A. Rigney
North Carolina State College
Raleigh, North Carolina

Drs. W. W. Schottstaedt and James Hagans
University of Oklahoma
Oklahoma City 4, Oklahoma

Dr. Malcolm E. Turner, Jr.
Medical College of Virginia
Richmond, Virginia

Dr. Colin White
Yale School of Medicine
New Haven, Connecticut

Dr. J. Yerushalmy
University of California
Berkeley 4, California

For those unable to train during the academic year, an unusual opportunity is provided by a cooperative GRADUATE SESSION OF STATISTICS IN THE HEALTH SCIENCES sponsored by these Program Directors and made possible by a training grant from the PHS, NIH. For information concerning available stipends and course offerings at elementary, intermediate, or advanced levels for the summers of 1961 and 1962, write Dr. Jacob E. Bearman, University of Minnesota, Minneapolis, Minnesota.

FELLOWSHIPS FOR MATHEMATICS GRADUATES

The Division of Preventive Medicine and the Institute of Industrial Health of the College of Medicine, The University of Cincinnati, with the support of the Public Health Service, are instituting a training program in Epidemiology and Biostatistics. As part of this program training and experience will be given in biology and research to bright, young, and prospective candidates for a Ph.D. degree in Mathematics or Statistics. The prospective fellow should have his B.A. or B.S. in Mathematics. He should have decided on a department at which he will want to complete his graduate work. Above all he must have demonstrated a level of achievement that will make him an acceptable candidate for an advanced degree at any university.

The student will be exposed also to research projects; working with medical and other fellows in this program and in the Institute of Industrial Health. He will be expected to write a publishable paper in some aspects of Mathematical Biology at the end of his two year stay. Arrangements may also be made for him to take additional work or gain additional research experience at another institution. The candidate will receive invaluable experience for a future career in the area of Biology or Mathematics and Statistics. The stipend will depend on qualifications and marital status and will range from $3400 to $4400 plus tuition.

For further information write to Dr. Theodor D. Sterling, Department of Preventive Medicine, College of Medicine, University of Cincinnati, Cincinnati, Ohio.

ON-THE-JOB TRAINEE APPOINTMENTS

The Department of Statistics at Roswell Park Memorial Institute has been awarded a grant by the National Cancer Institute and in conjunction with the University of Buffalo offers a two-year graduate on-the-job training program leading to a Master of Science degree in biostatistics.

Students with a Bachelor's degree, and possessing a good background in mathematics (through calculus), and course work in physical, biological or behavioral sciences who will begin their graduate work next fall are sought as applicants for these appointments. In addition to taking graduate course work, trainees will receive on-the-job training in research studies that are currently in progress at Roswell Park Memorial Institute. Trainees will receive an average annual stipend of $2300.00 plus tuition.

Requests for further information and application forms should be addressed to Miss Barbara A. Cote, c/o Dr. Irwin D. J. Bross, Chief, Department of Statistics, Roswell Park Memorial Institute, 666 Elm Street. Buffalo 3, New York.

BIOMETRIE—PRAXIMETRIE

Tome I, N° 3-4 *Juillet-Octobre, 1960*

Editorial

Utilisation de calculateurs électroniques dans l'analyse d'expériences avec répétitions et de groupes d'expériences de même plan expérimental. F. YATES

Planification des expériences d'alimentation des vaches laitières... H. L. LUCAS

La sélection artificielle suivant des caractères quantitatifs..... H. L. LE ROY

La répartition des sexes chez des cucurbitacées hybrides et 1 leur relation avec une distribution contagieuse de Neyman..... F. WEILING

La puissance du critère F dans l'analyse de la variance de plans en blocs au hasard, nomogrammes pour le choix du nombre de répétitions.... M. KEULS

Développement des méthodes biométriques et statistiques dans la recherche agronomique au Congo Belge et au Ruanda-Urundi..... J. -M. HENRY

Développement des méthodes biométriques et statistiques dans la recherche agronomique en Belgique..... L. MARTIN

L'emploi des méthodes biométriques dans l'étude et l'expérimentation des variétés de betterave sucrière en Belgique.... M. SIMON ET N. ROUSSEL

Les méthodes d'échantillonnage en sylviculture..... ANNE LENGER

Radiosensibilité des graines d'andropogon issues de terrains uranifères et non-uranifères du Katanga..... DR. J. MEWISSEN, J. DAMBLON, ET Z. M. BACQ

Dix années d'activités de la Biometric Society en Belgique, au Congo Belge et au Ruanda-Urundi

Revue bibliographique.

Tome II, N° 1 *Mars, 1961*

Quelques problèmes de numérations bactériennes..... H. GRIMM

Comparaison de deux méthodes d'analyse de données non orthogonales dans le cas d'expériences à deux facteurs..... P. GILBERT ET B. GROSSMANN

Analyse de données non orthogonales dans le cas d'une expériences à deux facteurs..... A. ROTTI

Revue bibliographique: D. J. FINNEY: Experimental Design and its Statistical Basis..... L. MARTIN

Publishers for the

SOCIETE ADOLPHE QUETELET
7 rue Héger-Bordet, Brussels 1, Belgium

Presses Universitaires de Bruxelles
50, av. Franklin Roosevelt
Brussels

JOURNAL OF THE
AMERICAN STATISTICAL
ASSOCIATION

Volume 56 June 1961 Number 294

TABLE OF CONTENTS

Confidence Curves: An Omnibus Technique for Estimation and Testing Statistical Hypotheses ..... ALLAN BIRNBAUM
Changes in the Size Distribution of Dividend Income ..... EDWIN B. COX
Note on Curve Fitting with Minimum Deviations by Linear Programming ..... WALTER D. FISHER
Bivariate Logistic Distributions ..... E. J. GUMBEL
Partial Correlations in Regression Computations .... ROBERT L. GUSTAFSON
An Analysis of Consistency of Response in Household Surveys
CAROL M. JAEGER AND JEAN L. PENNOCK
Multiple Regression Analysis of a Poisson Process .... DALE W. JORGENSON
Factorial Treatments in Rectangular Lattice Designs
CLYDE Y. KRAMER AND LEROY S. BRENNA
Significance Tests in Discrete Distributions ..... H. O. LANCASTER
Exact and Approximate Distributions for the Wilcoxon Statistic with Ties ..... SHIRLEY Y. LEHMAN
The Use of Sample Quasi-Ranges in Setting Confidence Intervals for the Population Standard Deviation
FRED C. LEONE, C. W. TOPP, AND Y. H. RUTENBERG
Randomized Rounded-Off Multipliers in Sampling Theory
M. N. MURTHY AND V. K. SETHI
Unbiased Componentwise Ratio Estimation
D. S. ROBSON AND CHITRA VITHAYASAL
A Note on Measurement Errors and Detecting Real Differences ..... EUGENE ROGOT
A Quarterly Econometric Model of the United States
LOWELL E. GALLAWAY AND PAUL E. SMITH
On the Use of Partially Ordered Observations in Measuring the Support for a Complete Order ..... R. F. TATE
The Statistical Work of Oskar Anderson ..... GERHARD TINTNER
A Problem Concerned with Weighting of Distributions ..... COLERIDGE A. WILKINS

For further information, please contact:
AMERICAN STATISTICAL ASSOCIATION
1757 K Street, N.W., Washington 6, D. C.

TECHNOMETRICS

A Journal of Statistics for the Physical, Chemical and Engineering Sciences

Vol. 3, No. 1 February, 1961

CONTENTS

Cumulative Sum Charts.....E. S. PAGE
Average Run Lengths in Cumulative Chart Quality Control Schemes
P. L. GOLDSMITH AND H. WHITFIELD
Prediction Regions for Several Predictions from a Single Regression Line
GERALD J. LIEBERMAN
The Robustness of Certain Life Testing Procedures Derived from the Exponential Distribution.....MARVIN ZELEN AND MARY C. DANNEMILLER
An Application of a Balanced Incomplete Block Design..PETER W. M. JOHN
Multi-component Systems and Structures and Their Reliability
Z. W. BIRNBAUM, J. D. ESARY AND S. C. SAUNDERS
An Asymptotic Distribution for an Occupancy Problem with Statistical Applications.....M. HALPERIN AND G. L. BURROWS
Outliers in Patterned Experiments: A Strategic Appraisal
IRWIN D. J. BROSS
Book Reviews.....ACHESON J. DUNCAN AND WILLIAM BERANEK
Statistical Programs for High Speed Computers
Notices

Vol. 3, No. 2 May, 1961

CONTENTS

General Considerations in the Analysis of Spectra.....G. M. JENKINS
Mathematical Considerations in the Estimation of Spectra
EMANUEL PARZEN
Discussion, Emphasizing the Connection Between Analysis of Variance and Spectrum Analysis.....JOHN W. TUKEY
Some Comments on Spectral Analysis of Time Series.....N. R. GOODMAN
Comments on the Discussions of Messrs. Tukey and Goodman
G. M. JENKINS AND EMANUEL PARZEN
Spectral Analysis Combining a Bartlett Window with an Associated Inner Window.....THOMAS H. WONNACOTT
Frequency Response from Stationary Noise: Two Case Histories
N. R. GOODMAN, S. KATZ, B. H. KRAMER AND M. T. KUO
The Modified Gauss-Newton Method for the Fitting of Non-Linear Regression Functions by Least Squares.....H. O. HARTLEY
On the Possibility of Improving the Mean Useful Life of Items by Eliminating Those with Short Lives.....G. S. WATSON AND W. R. WELLS
The Optimum Allocation of Spare Components in System
DONALD F. MORRISON
Use of Tables of Percentage Points of Range and Studentized Range
H. LEON HARTER
Book Reviews.....C. DANIEL
Statistical Programs for High Speed Computers
Notices

Technometrics is published quarterly in February, May, August, and November. The annual non-member subscription rate is $8.00. To members of the American Statistical Association and the American Society for Quality Control the rate is $6.00. Checks should be made payable to *Technometrics* and addressed to *Technometrics*, Post Office Box 587, Benjamin Franklin Station, Washington 6, D. C.

INTERNATIONAL JOURNAL OF ABSTRACTS STATISTICAL THEORY AND METHOD

A Journal of the International Statistical Institute

The aim of this journal of abstracts is to give **complete coverage of published papers in the field of statistical theory** (including associated aspects of probability and other mathematical methods) **and new published contributions to statistical method.**

All contributions in the following five journals—being wholly devoted to this field—are abstracted: *Annals of Mathematical Statistics; Biometrika; Journal, Royal Statistical Society (Series B); Bulletin of Mathematical Statistics; Annals, Institute of Statistical Mathematics;* and a further group of six journals are abstracted on a **virtually complete** basis as follows: *Biometrics; Metrika; Metron; Review, International Statistical Institute; Technometrics; Sankhyā.* There are about 250 other journals partly devoted to statistical theory and method from which the appropriate papers are abstracted.

The abstracts are about **400** words long—the recommendation of UNESCO for the "long" abstract service: they are in the English language although the original language of the paper is noted on the abstract together with the name of abstractor. In addition to the address of the author(s) are given in detail to facilitate contact in order to obtain further detail or request an off-print. **The journal is published quarterly and contains approximately 1000 abstracts per year.**

A scheme of classification has been developed for the abstracts that is flexible and facilitates the transfer of code numbers to punched cards. **A unique aspect of this journal** is that the pages are colour-tinted according to the main sections of classification. This method of colour-coding the pages provides a distinctive and powerful visual aid in the identification of abstracts in whatever manner the journal is filed for reference.

Annual Subscription	£5 (U.S.A. and Canada $16.00)
Single Number	30s (U.S.A. and Canada $4.50)

OLIVER AND BOYD LTD.
Tweeddale Court, 14 High Street, Edinburg, 1

INFORMATION FOR CONTRIBUTORS

Manuscripts

Contributions for *Biometrics* may be addressed to Dr. Ralph A. Bradley, Department of Statistics, The Florida State University, Tallahassee, Florida, U.S.A.; authors residing in the following Society Regions can expedite consideration of papers by submitting them to the appropriate Associate Editor, namely; BRITISH REGION: Dr. S. C. Pearce, East Malling Research Station, East Malling, Maidstone, Kent, England; AUSTRALASIAN REGION: Dr. E. A. Cornish, University of Adelaide, Adelaide, Australia; FRENCH REGION: Dr. Georges Teissier, Faculté des Sciences de Paris, 1 rue V. Cousin, Paris, France. QUERIES, NOTES, and related correspondence should be directed to Dr. D. J. Finney, Department of Statistics, University of Aberdeen, Meston Walk, Old Aberdeen, Scotland. Books and material for Book Reviews should be sent to Mr. J. G. Skellam, The Nature Conservancy, 19 Belgrave Square, London, S.W. 1, England.

MANUSCRIPTS must be submitted in *triplicate*, with typescript doublespaced throughout. Marginal notes may obviate typographical difficulties presented for complicated formulae or tables—authors should not attempt editorial instructions or markings for the printer. TABLES should be identified by arabic number and by a short descriptive title. ILLUSTRATIONS should also be identified by arabic number and by a brief caption. (Captions should not be included in illustrations, but should be typewritten collectively on an accompanying sheet.) Originals should be approximately 8.5 x 11 in. (21.5 x 28 cm.). The original of each *chart*, *diagram*, or *graph* should be executed in black on white drawing paper or board, on blue tracing linen, or on coordinate paper ruled in blue only; coordinate lines to be reproduced should be ruled in black. For printing, illustrations may be reduced to ½ or ⅓ original dimensions. Lines should therefore be of sufficient thickness, and decimal points, periods, and stippled dots should be solid black circles large enough to reproduce well. Lettering and numerals should be at least 1 mm. high when reproduced in a cut 3 in. (7.5 cm.) wide. Photographs should be prints on glossy paper with strong contrasts, and if grouped in a plate should be mounted contiguously. All tables and illustrations should be mentioned explicitly in the text. REFERENCES (BIBLIOGRAPHIC) should be collectively listed alphabetically by author; textual citation by author and year is preferred.

Abstracts

Abstracts of papers presented at meetings of the Biometric Society or of its regions are printed in *Biometrics* following such meetings. They should be submitted to the person designated to receive them for a particular meeting in exactly the form published in *Biometrics* (except for an Abstract Number), doublespaced on bond paper, and in duplicate. Use of formulae requiring display printing is to be avoided.

Notices, Announcements, and Biometric Society Reports

International and regional reports and notices should be submitted by the appropriate officers of the Society and its Regions in *duplicate* doublespaced on separate sheets exactly as they are to be printed in *Biometrics*. Other material to be printed in News and Announcements should also be submitted doublespaced and in duplicate.

Sustaining Members of the Biometric Society

Abbott Laboratories
American Cancer Society, Inc.
General Foods Corporation, Research Center
Merck, Sharp and Dohme Research Laboratories
Schering Corporation
Smith, Kline and French Laboratories
E. R. Squibb and Sons
The Upjohn Company
Wallace Laboratories, Division of Carter Products
Wyeth Institute of Applied Biochemistry

BACK ISSUES

Back issues of *Biometrics* are available at the following postage-paid prices in U.S.A. currency:

Year	Volume	Number	Price per Single Number	Price per Volume (unbound)
1945	1	1 to 6	$1.00	$6.00
1946	2	1 to 6	1.00	6.00
1947	3	1 to 4	1.50	5.00
1948	4	1 to 4	1.50	5.00
1949	5	1 to 4	1.50	5.00
1950	6	1 to 4	1.50	5.00
1951	7	1 to 4	2.00	8.00
1952	8	1 to 4	2.00	8.00
1953	9	1 to 4	2.00	8.00
1954	10	1 to 4	2.00	8.00
1955	11	1 to 4	2.00	8.00
1956	12	1 to 4	2.00	8.00
1957	13	1 to 4	2.00	8.00
1958	14	1 to 4	2.00	8.00
1959	15	1 to 4	2.00	8.00
1960	16	1 to 4	2.00	8.00

Reprints of individual articles are not available except to authors at the time of printing. Three special issues are among the numbers listed above. They are:

1947	Volume 3	Number 1	The Analysis of Variance
1951	Volume 7	Number 1	Components of Variance
1957	Volume 13	Number 3	The Analysis of Covariance

Also available are:

Fishery Reprint Series (Selected reprints from Vol. 5)	$1.00
Subject Index (Volumes 1–10)	1.00
Proceedings, International Biometric Symposium, Campinas, Brazil, 1955.	1.00

Inquiries, non-member subscriptions, and orders for back issues and other material listed above should be addressed to: BIOMETRICS, DEPARTMENT OF STATISTICS, THE FLORIDA STATE UNIVERSITY, TALLAHASSEE, FLORIDA, U.S.A.